Математический анализ
Семестр 2.
Сводка результатов

Оглавление

1. Нормированные пространства

Нормированное пространство — это векторное пространство с метрикой. При этом метрика и структура векторного пространства должны быть согласованы (метрика инвариантна относительно сдвигов, и при умножении вектора на число его длина умножается на модуль этого числа).

Формально говоря, пусть для каждого вектора x в векторном пространстве X определено действительное число $\|x\|$, его *норма*, причём:

• норма неотрицательна и равна нулю только для нулевого вектора;

• $\|x + y\| \leqslant \|x\| + \|y\|$ для любых векторов $x, y \leqslant X$ (*неравенство треугольника*);

•$\|\lambda x\| = |\lambda| \cdot \|x\|$ для любого вектора x и для любого числа λ.

Тогда множество X становится *нормированным векторным пространством.*

Мы будем рассматривать нормированные пространства над полем $\mathbb{R}$ (λ вещественно), однако аналогичное определение даётся и для $\mathbb{C}$, тогда $|\lambda|$ есть модуль комплексного числа λ.

Расстояние между точками x, y в нормированном пространстве определяется как $\|x-y\|$. Оно удовлетворяет аксиомам метрического пространства.

Примеры. Норму вектора $(x_1, \ldots, x_n) \in \mathbb{R}^n$ можно определить как $|x_1| + \ldots + |x_n|$, как $\sqrt{|x_1|^2 + \ldots + |x_n|^2}$ или как $\max(|x_1|, \ldots, |x_n|)$. Получаются три различные нормы, которые обозначают $\|x\|_1$, $\|x\|_2$ и $\|x\|_\infty$. Обобщение: положим

$$\|x\|_p = (x_1^p + \ldots + x_n^p)^{1/p}.$$

(при $p = \infty$ надо перейти к пределу).

1 Проверьте, что при $p > 1$ это определение задаёт норму, а при $p < 1$ — нет. (В трёх частных случаях $p = 1$, $p = 2$ и $p = \infty$ проверка несколько проще, чем в общем случае.)

Можно определить норму в произведении двух нормированных пространств X и Y, положив $\|(x, y)\|_{X \times Y}$ равным $\|x\|_X + \|y\|_Y$. Этой формуле соответствует $\|\cdot\|_1$-норма; аналогично можно определить аналог нормы $\|\cdot\|_p$.

Пусть в одном и том же векторном пространстве заданы две нормы. Им соответствуют две метрики. Эти метрики эквивалентны тогда и только тогда, когда нормы отличаются не более чем на константу: $(1/C)\|x\|_1 \leqslant \|x\|_2 \leqslant C\|x\|_1$ для некоторого $C > 0$ и для всех x. Рассмотренные выше метрики в $\mathbb{R}^n$ эквивалентны. Более того, *любые две нормы в конечномерном векторном пространстве эквивалентны.*

В самом деле, достаточно доказать, что в $\mathbb{R}$ любая норма $\|\cdot\|$ эквивалентна норме $\|\cdot\|_1$ (сумме модулей координат). Рассмотрим стандартный базис $e_1, \ldots, e_n$ в $\mathbb{R}^n$. Тогда

$$\|(x_1, \ldots, x_n)\| = \|x_1 e_1 + \ldots + x_n e_n\| \leqslant \\ \leqslant \|x_1 e_1\| + \ldots + \|x_n e_n\| = |x_1|\|e_1\| + \ldots + |x_n|\|e_n\| \leqslant C\|x\|_1$$

где $C = \max(\|e_1\|, \ldots, \|e_n\|)$. В одну сторону оценка получена (новая норма не сильно больше старой). Отсюда (и из неравенства треугольника для новой нормы) следует, что новая норма является непрерывной функцией $\mathbb{R}^n \to \mathbb{R}$ (относительно старой нормы в $\mathbb{R}^n$). На единичной сфере (в смысле старой нормы) эта функция всюду положительна. Поэтому в силу её непрерывности и компактности сферы эта функция не только ограничена на сфере (эта оценка у нас и так есть), но и отделена от нуля некоторым $\varepsilon > 0$, откуда $\|x\| \geqslant \varepsilon\|x\|_1$ (на единичной сфере, а потому и везде).

Какие бывают нормы, скажем, на плоскости? Геометрически это можно понять так. Каждая норма задаёт единичный шар (векторы с нормой не больше 1). Зная шар, можно восстановить норму однозначно (норма вектора — наименьшее число, на которое надо поделить вектор, чтоб он попал в шар). Шар — выпуклое центрально симметричное множество, и с каждой прямой, проходящей через нуль (=с каждым одномерным подпространством), он пересекается по отрезку ненулевой (но конечной длины). И это всё — любое

выпуклое центрально симметричное множество, пересекающееся с каждым одномерным подпространством по отрезку конечной длины, есть единичный шар некоторой нормы.

2 Нарисовать шары для рассмотренных норм в $\mathbb{R}^2$ и придумать какую-нибудь норму, с ними не совпадающую.

Пусть X, Y — два нормированных пространства, $A\colon X \to Y$ — линейное отображение. Тогда следующие свойства равносильны:

- A непрерывно всюду;
- A непрерывно в нуле (пространства X);
- A непрерывно хотя бы в одной точке;
- A ограничено на единичном шаре (найдётся такое C, что $\|Ax\| \leqslant C$ при любом x с $\|x\| \leqslant 1$);
- найдётся такое C, что $\|Ax\| \leqslant C\|x\|$ при всех $x \in X$.

3 Проверьте, что эти свойства равносильны (и потому непрерывные линейные операторы называют также *ограниченными*, имея в виду ограниченность на единичном шаре).

Наименьшее из чисел C, для которых выполняются указанные требования, называется *нормой* оператора A. Другими словами,

$$\|Ax\| \leqslant \|A\| \cdot \|x\|,$$

и $\|A\|$ нельзя заменить на меньшее число.

4 Проверьте, что действительно получается норма на линейном пространстве непрерывных операторов из X в Y.

Это пространство обозначают $L(X, Y)$.

5 Для конечномерных нормированных пространств E и F любой линейный оператор $E \to F$ непрерывен.

6 Пусть $X = \mathbb{R}^n$, $Y = \mathbb{R}$. Тогда $L(X, Y)$ канонически изоморфно $\mathbb{R}^n$. Поэтому каждой норме в $\mathbb{R}^n$ соответствует некоторая другая норма в $\mathbb{R}^n = L(\mathbb{R}^n, \mathbb{R})$. Проверьте, что норме $\|\cdot\|_1$ соответствует $\|\cdot\|_\infty$ и наоборот, а также что евклидовой норме $\|\cdot\|_2$ соответствует она сама.

7 (Продолжение) Пусть $p, q > 1$ и $1/p + 1/q = 1$. Проверьте, что норме $\|\cdot\|_p$ соответствует норма $\|\cdot\|_q$.

8 Найдите норму линейного оператора $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ с матрицей

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

если в $\mathbb{R}^2$ (в обоих случаях) рассматривается норма (**а**) $\|\cdot\|_1$; (**б**) $\|\cdot\|_2$; (**в**) $\|\cdot\|_\infty$.

Нормированные пространства, являющиеся полными (всякая фундаментальная последовательность имеет предел), называют *банаховыми*.

Пространство $L(X, Y)$ является полным, если полно пространство Y.

2. Общее определение производной и её свойства

Пусть функция f определена в окрестности точки a в нормированном пространстве E и принимает значения в нормированном пространстве F. Говорят, что f *дифференцируема* в точке a, если существует ограниченный линейный оператор $L : E \to F$, для которого

$$f(a+h) = f(a) + Lh + o(\|h\|)$$

при $h \to 0$. Формально это означает, что для всякого $\varepsilon > 0$ существует $\delta > 0$, при котором

$$\|f(a+h) - f(a) - Lh\|_F \leqslant \varepsilon \|h\|_E$$

для любого $h \in E$ с $\|h\|_E < \delta$.

Если $E = \mathbb{R}$, то дифференцируемость функции в точке a равносильна наличию предела

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

В самом деле, линейный оператор $\mathbb{R} \to F$ есть умножение на вектор, то есть отображение $h \mapsto hf$ для некоторого $f \in F$, а требование на $o(h)$ после деления на h превращается в определение предела.

Дифференцируемая в точке функция непрерывна в этой точке (здесь существенно, что производная есть *ограниченный* оператор).

Сумма двух дифференцируемых функций, определённых в окрестности точки $a \in E$ и принимающих значения в нормированном пространстве F, дифференцируема, и производная суммы равна сумме производных. (Аналогично для разности и для умножения на число.)

Линейный оператор $L : E \to F$ дифференцируем в любой точке и производная в любой точке равна самому L.

Функция f из E в $F_1 \times F_2$ дифференцируема тогда и только тогда, когда её компоненты $f_1 : E \to F_1$ и $f_2 : E \to F_2$ дифференцируемы, и

$$f'(a)h = (f_1'(a)h, f_2'(a)h).$$

Для дифференцирования «произведения» нужно предполагать непрерывность операции произведения. Пусть E, F, G — векторные пространства. Напомним, что отображение

$$B\colon E \times F \to G$$

называется *билинейным*, если отображение $f \mapsto B(e,f)$ является линейным для любого $e \in E$, а отображение $e \mapsto B(e,f)$ является линейным для любого $f \in F$. Билинейные отображения находятся во взаимно однозначном соответствии с линейными отображениями из E в векторное пространство линейных отображений $F \to G$ (а также с линейными отображениями из F в векторное пространство линейных отображений $E \to G$).

Если E, F, G являются нормированными пространствами, то среди билинейных операторов выделяются непрерывные. Непрерывность в любой точке равносильна непрерывности в нуле, ограниченности в окрестности нуля, а также существованию константы C, при которой

$$\|B(e,f)\|_G \leqslant C\|e\|_E \cdot \|f\|_F$$

при всех $e \in E$ и $f \in F$. Наименьшая из таких констант C называется *нормой* непрерывного билинейного отображения. Непрерывные билинейные отображения $E \times F \to G$ находятся во взаимно однозначном соответствии с $L(E, L(F, G))$, это соответствие сохраняет норму. (Аналогично для $L(F, L(E, G))$.)

После такой подготовки (излишней в конечномерном случае, когда всё линейное непрерывно) можно сформулировать правило дифференцирования произведения. Пусть функции f и g определены в окрестности точки a нормированного пространства E и принимают значения в нормированных пространствах F и G. Пусть $B\colon F \times G \to U$ — непрерывное билинейное отображение со значениями в нормированном пространстве U. Если функции f и g дифференцируемы в точке a, то функция

$$s\colon x \mapsto B(f(x), g(x))$$

дифференцируема в точке a, и

$$s'(a)h = B(f'(a)h, g(a)) + B(f(a), g'(a)h).$$

Доказательство повторяет одномерное:

$$\begin{aligned} s(a+h) &= B(f(a+h), g(a+h)) = \\ &= B(f(a) + f'(a)h + o(h), g(a) + g'(a)h + o(h)) = \\ &= B(f(a), g(a)) + B(f'(a)h, g(a)) + B(f(a), g'(a)h) + o(h). \end{aligned}$$

При этом мы пользовались линейностью B по каждому аргументу, тем, что $B(f'(a)h, g'(a)h) = o(h)$ и тем, что $B(o(h), O(1)) = B(O(1), o(h)) = o(h)$. Два последних утверждения опираются на непрерывность B.

9 Найти производную непрерывной билинейной функции $B\colon E \times F \to U$ как отображения нормированных пространств. В случае $E = F$ найти производную квадратичной функции $x \mapsto B(x, x)$. [Напоминание: билинейная форма B не обязана быть симметричной.]

Аналогом функции $x \mapsto 1/x$ является отображение обращения оператора. Пусть X — произвольное банахово пространство. Оператор $A \in L(X, X)$ *обратим*, если существует оператор $B \in L(X, X)$, для которого $AB = E$ и $BA = E$ (где $E \in L(X, X)$ — тождественный оператор ($Ex = x$ при всех $x \in X$). (В бесконечномерном случае важны оба условия $AB = E$ и $BA = E$.)

Множество $U \subset L(X, X)$ обратимых операторов открыто и функция обращения $i\colon A \mapsto A^{-1}$ дифференцируема в любой точке множества U. При этом

$$i'(A)H = -i(A) \circ H \circ i(A)$$

Другими словами,

$$(A + H)^{-1} = A^{-1} - A^{-1} \circ H \circ A^{-1} + o(\|H\|)$$

для любого обратимого оператора A при $H \to 0$.

(Таким образом, в обычной формуле для $d(1/x) = -dx/x^2$ нужно лишь правильно выбрать порядок действий в правой части: $d(1/x) = -(1/x)\,dx(1/x)$.)

Чтобы убедиться в этом, сначала рассмотрим случай $A = E$ (единичный оператор).

Лемма: если $\|H\| < 1$, то

$$(E + H)^{-1} = E - H + H^2 - H^3 + \ldots$$

Ряд в правой части абсолютно сходится, поскольку $\|H^n\| \leqslant \|H\|^n$ и $\|H\| < 1$. Отсюда же следует, что все члены ряда после первого в сумме дают $o(\|H\|)$ при $H \to 0$. Так что операция обращения определена в окрестности единицы и имеет там указанную выше производную.

В произвольной точке:

$$(A + H)^{-1} = [A(E + A^{-1}H)]^{-1} = (E + A^{-1}H)^{-1}A^{-1} = \\ = (E - A^{-1}H + o(A^{-1}H))A^{-1} = A^{-1} - A^{-1}HA^{-1} + o(\|H\|),$$

что и требуется.

10 Докажите аналогичную формулу для операции обращения $L(X, Y) \to L(Y, X)$ для произвольных банаховых пространств X, Y.

Наконец, для отображений нормированных пространств имеется теорема о *производной сложной функции*. Пусть функция f определена в окрестности точки $a \in X$ и принимает значения в пространстве Y, а функция g определена в окрестности точки $f(a)$ и принимает значения в пространстве Z. Пусть f дифференцируема в a, а g дифференцируема в $f(a)$. Тогда композиция («сложная функция») $g \circ f$ определена в окрестности точки a и дифференцируема в a, причём

$$(g \circ f)'(a) = g'(f(a)) \circ f'(a)$$

(в правой части стоит композиция линейных операторов).

Доказательство повторяет одномерное: в силу непрерывности f в точке a значение $f(a+h)$ при малых h попадает в область определения функции g. При этом

$$\begin{aligned}(g \circ f)(a+h) = g(f(a+h)) = g(f(a) + [f'(a)h + o(h)]) = \\ = g(f(a)) + g'(f(a))k + o(\|k\|) = g(f(a)) + g'(f(a))f'(a)h + o(h);\end{aligned}$$

через k обозначено содержимое квадратной скобки; мы пользуемся тем, что $k = O(\|h\|) + o(\|h\|) = O(\|h\|)$ и $o(\|k\|) = o(O(\|h\|)) = o(\|h\|)$.

11 Пусть функция $f\colon \mathbb{R} \to E$ принимает значения на единичной сфере в евклидовом пространстве E. Тогда вектор скорости $f'(t)$ перпендикулярен $f(t)$.

3. Частные производные

Пусть функция f определена в окрестности точки a нормированного пространства X и принимает значения в пространстве Y. Для любого вектора $v \in X$ можно рассмотреть функцию

$$t \mapsto f(a + tv),$$

определенную в окрестности нуля в $\mathbb{R}$ и принимающую значения в пространстве Y. Если эта функция имеет производную в нуле, то она называется *производной вдоль вектора v функции f в точке a*.

Если функция f дифференцируема в точке a, то производная вдоль вектора v равна $f'(a)v$ по теореме о производной композиции.

12 Приведите пример функции $f\colon \mathbb{R}^2 \to \mathbb{R}$, которая имеет в нуле производные вдоль любого вектора, но не дифференцируема. [Указание: одна

из возможных причин — производная вдоль v зависит от v нелинейно, но возможны и другие причины.]

13 (Продолжение) Приведите чуть более сложный пример, когда производная вдоль любого вектора есть в любой точке, но функция тем не менее в некоторой точке не дифференцируема.

[Может ли она быть не дифференцируема при этом ни в одной точке? Наверно, это должно быть известно.]

Понятие производной вдоль вектора объясняет смысл величины $f'(a)v$: это скорость образа точки, когда сама точка проходит через a со скоростью v. (Такая интерпретация полезна при дифференцировании функций на многообразиях.)

Для функции f, определённой в окрестности точки $(a_1, \ldots, a_n) \in \mathbb{R}^n$, можно рассмотреть её производные вдоль базисных векторов $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$ и т.д. Из определения ясно, что это будут производные функции f как функции одного из аргументов при фиксированных остальных аргументах. Скажем, производная вдоль e_1 в точке a есть производная функции

$$x \mapsto f(x, a_2, a_3, \ldots, a_n)$$

в точке a_1. Её обозначают обычно

$$\frac{\partial f}{\partial x_1} f(a_1, \ldots, a_n)$$

и называют *частной производной* функции f по первому аргументу в точке $(a_1, \ldots, a_n)$. (Обозначение $\partial f / \partial x_1$ не вполне корректно, поскольку неявно предполагает, что первый аргумент функции f будет всегда обозначаться x_1, но стандартно.)

Как мы уже говорили, само по себе существование частных производных (и даже производной вдоль всех векторов во всех точках) не гарантирует дифференцируемости, но если они непрерывны, то функция дифференцируема (как мы увидим).

Частные производные однозначно определяют производную (линейный функционал однозначно определяется своими значениями на базисных векторах).

Чтобы избежать путаницы и отличить производную от частных производных, её называют также *полной производной* ил *полным дифференциалом.*

Функцию $f\colon \mathbb{R}^n \to \mathbb{R}^m$ можно рассматривать как набор из m функций

$$f_1, \ldots, f_m \colon \mathbb{R}^n \to \mathbb{R}.$$

Дифференцируемость функции f равносильна дифференцируемости каждой из f_i, которая в свою очередь влечёт существование частных производных. Эти частные производные (для данной точки a в $\mathbb{R}^n$) образуют матрицу с элементами $a_{ij} = \partial f_i / \partial x_j$, которая является матрицей линейного оператора $f'(a)$. Таким образом,

$$[f'(a)(h_1, \ldots, h_n)]_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} h_j$$

(квадратные скобки в левой части обозначают i-ю компоненту вектора в $\mathbb{R}^m$).

Матрица линейного оператора $f'(a)$ называется *матрицей Якоби*. В случае $m = n$ можно рассмотреть её определитель, который называется *якобианом*; его геометрический смысл — коэффициент, на который умножается объём бесконечно малого участка около a (это, конечно, требует уточнения и доказательства).

Можно говорить о частных производных и в более общем случае. Пусть имеется отображение произведения банаховых пространств $X_1 \times \ldots \times X_n$ в произведение $Y_1 \times \ldots \times Y_m$. Тогда можно рассмотреть j-ую компоненту этого отображения (при $1 \leqslant j \leqslant m$) и фиксировать в ней все аргументы, кроме i-го. Получится функция из X_i в Y_j, производную которой можно назвать частной производной j-ой компоненты по i-ой переменной. Из таких производных можно образовать матрицу, но уже не числовую, а из линейных операторов, и умножать эту матрицу на столбец по обычному правилу.

4. Теорема о конечном приращении

Для дифференцируемых функций $f\colon \mathbb{R} \to X$ при многомерном X нет теоремы Лагранжа: $f(b) - f(a)$ может не равняться $f'(\xi)(b-a)$ ни для какой точки $\xi \in (a, b)$.

(Простейший пример: машина проезжает по кольцевой дороге, возвращаясь в исходную точку, но скорость ни в какой момент не обращается в нуль.)

Тем не менее по-прежнему средняя скорость не превосходит максимальной мгновенной: *если функция $f\colon [a,b] \to X$ непрерывна на $[a,b]$ и дифференцируема на (a,b), причём $\|f'(\xi)\| \leqslant M$ при любом $\xi \in (a,b)$, то $\|f(b) - f(a)\| \leqslant M(b-a)$.*

Переформулировка:

$$\|f(b) - f(a)\| \leqslant \sup_{\xi \in (a,b)} \|f'(\xi)\| \cdot (b-a)$$

При этом производную можно рассматривать как вектор в том же пространстве X (поскольку $L(\mathbb{R}, X)$ как нормированное пространство отождествляется с X).

Для доказательства достаточно убедиться, что

$$\|f(b) - f(a)\| \leqslant M'(b - a)$$

для произвольного $M' > M$. Фиксируем такое M' и назовём отрезок $[\alpha, \beta] \subset [a, b]$ плохим, если для него это неравенство (с данным M') нарушается. У плохого отрезка плоха одна из половин, делим так и дальше и берём общую точку вложенных плохих отрезков. В ней производная по модулю не больше M, поэтому малые отрезки не могут быть плохими. (Это рассуждение не годится, если общая точка оказывается с краю, но можно немного сдвинуть a и b, оставляя отрезок плохим.)

Другое доказательство для евклидовых пространств: спроектируем f на единичный вектор в направлении $f(b) - f(a)$; производная проекции равна проекции производной и не превосходит M по модулю, осталось сослаться на одномерную теорему.

Для случая отображения $X \to Y$ при многомерном X: *если функция $f\colon X \to Y$ определена и дифференцируема на открытом множестве, содержащем точки $a, b \in X$ вместе с отрезком $[a, b] \subset X$, то*

$$\|f(b) - f(a)\| \leqslant \sup_{\xi \in (a,b)} \|f'(\xi)(b - a)\|.$$

В правой части линейный оператор $f'(\xi) \in L(X, Y)$ применяется к вектору $(b - a) \in X$, и получается вектор пространства Y.

Это утверждение сводится к одномерному для функции $t \mapsto f(a + t(b - a))$. Его можно ослабить, написав в правой части

$$\Big[\sup_{\xi \in (a,b)} \|f'(\xi)\|\Big] \cdot \|b - a\|.$$

Следствие: если функция дифференцируема на открытом множестве и производная равна нулю во всех его точках, то эта функция локально постоянна.

14 Докажите, что если функция определена и дифференцируема на выпуклом множестве, причём производная ограничена, то она равномерно непрерывна.

Иногда полезна и такая форма теоремы о конечном приращении: если $A \in L(X, Y)$ — произвольный ограниченный линейный оператор, то

$$\|f(b) - f(a) - L(b - a)\| \leqslant \Big[\sup_{\xi \in (a,b)} \|f'(\xi) - L\|\Big] \cdot \|b - a\|.$$

Она показывает, что если производная почти постоянна, то функция близка к аффинной.

5. Дифференцируемость и частные производные

Если частные производные определены в окрестности некоторой точки и непрерывны в этой точке, то функция дифференцируема в этой точке.

Это утверждение верно для функций из $\mathbb{R}^n$ в $\mathbb{R}^m$, а также для случая произведений банаховых пространств.

Докажем его для произведения двух пространств. Пусть функция f определена в окрестности точки (a, b) пространства $A \times B$, во всей этой окрестности есть обе частные производные, причём одна из них (скажем, по B-координате) непрерывна в точке (a, b). (Достаточно непрерывности одной из двух частных производных.)

Проверим, что есть полная производная в точке (a, b). В самом деле, переход от (a, b) к $(a+\Delta a, b+\Delta b)$ можно разбить на две части — сначала сдвиг на Δa по первой координате, потом на Δb по второй. Для каждой части записываем изменение f и складываем:

$$\begin{aligned} f(a+\Delta a, b+\Delta b) = f(a, b) &+ \\ &+ [f(a+\Delta a, b) - f(a, b)] + \\ &\qquad + [f(a+\Delta a, b+\Delta b) - f(a+\Delta a, b)] \end{aligned}$$

Первая квадратная скобка есть $f'_A(a, b)\Delta a + o(\Delta a)$ непосредственно по определению частной производной. (Здесь f'_A обозначает частную производную вдоль A.) Вторую надо оценивать по теореме о конечном приращении:

$$\begin{aligned} \|f(a+\Delta a, b+\Delta b) - f(a+\Delta a, b) - f'_B(a, b)\Delta b\| \leqslant \\ \leqslant \sup_{\theta\in(0,1)} \|f'_B(a+\Delta a, b+\theta\cdot\Delta b) - f'_B(a, b)\| \cdot \|\Delta b\| \end{aligned}$$

В силу непрерывности частной производной f'_B первый множитель в последней строке есть $o(1)$ при $(\Delta a, \Delta b) \to 0$, что и требуется согласно определению производной.

Аналогичное рассуждение годится и для случая большего числа переменных; при этом нужна непрерывность всех частных производных, кроме одной.

6. Вторая производная

Пусть функция f определена в некотором открытом подмножестве U банахова пространства X и принимает значения в банаховом пространстве Y. Пусть она дифференцируема во всех точках множества U. Тогда её производная представляет собой отображение $U \to L(X, Y)$ (точке $x \in U$ ставится в соответствие оператор $f'(x) \in L(X, Y)$). Может оказаться, что это отображение дифференцируемо в точке $a \in U$. Тогда его производная называется *второй производной* функции f и обозначается $f''(a)$.

Согласно определению, $f''(a)$ является элементом пространства $L(X, L(X, Y))$. Это значит, что её можно применить к вектору $h \in X$, и получится оператор в $L(X, Y)$ (примерно равный $f'(a+h) - f'(a)$); в свою очередь этот оператор можно применить к ещё одному вектору $k \in X$. Таким образом, производной соответствует билинейное отображение

$$(h, k) \mapsto [f''(a)h]k$$

из $X \times X$ в Y. Оказывается, что *это отображение симметрично:*

$$[f''(a)h]k = [f''(a)k]h$$

(При этом не предполагается ничего, кроме самого факта существования второй производной; естественно, что в этом случае первая производная существует в некоторой окрестности точки a.)

Неформально говоря, причина этой симметрии — в симметричности второй разности. Пусть h — вектор пространства X, а f — функция, определённая на X. Тогда можно рассмотреть функцию

$$\Delta_h f\colon x \mapsto f(x+h) - f(h)$$

Оператор Δ_h можно считать аналогом дифференцирования вдоль вектора h. Так вот, операторы Δ_h и Δ_k для любых двух векторов h и k коммутируют:

$$\Delta_h \Delta_k f(a) = \Delta_k \Delta_h f(a) = f(a+h+k) - f(a+h) - f(a+k) + f(a).$$

Формальное доказательство симметрии второй производной основано на такой лемме:

$$\Delta_h \Delta_k f(a) = [f''(a)k]h + o((\|h\| + \|k\|)^2).$$

В самом деле, если эта лемма верна, то, поскольку левая часть её симметрична, то

$$[f''(a)h]k - [f''(x)k]h = o((\|h\| + \|k\|)^2).$$

Левая часть является билинейной функцией от h и k, поэтому такое возможно, лишь если она равна нулю. (Подставив th и tk вместо h и k для стремящегося к нулю коэффициента t, получим, что левая часть пропорциональна t^2, а правая есть $o(t^2)$, и потому коэффициент пропорциональности равен нулю.)

Осталось доказать лемму. Перепишем левую часть её как

$$\Delta_h \Delta_k f(a) = [\Delta_k f](a+h) - [\Delta_k f](a)$$

и применим теорему о конечном приращении к дифференцируемой функции $\Delta_k f$ и линейному оператору, являющемуся производной этой функции в точке a. Получим:

$$\|[\Delta_k f](a+h) - [\Delta_k f](a) - [\Delta_k f]'(a)h\| \leqslant \sup_{\theta\in[0,1]} \|[\Delta_k f]'(a+\theta h) - [\Delta_k f]'(a)\| \cdot \|h\| \qquad (*)$$

Далее заметим, что разность коммутирует с дифференцированием, так что в левой части $[\Delta_k f]'(a)$ равно $[\Delta_k (f')](a)$, что есть $f''(a)k + o(k)$ по определению второй производной. Поэтому

$$[\Delta_k f]'(a)h = [f''(a)k]h + o(k)h = [f''(a)k]h + o((\|k\| + \|h\|)^2),$$

и для доказательства леммы остаётся проверить, что правая часть неравенства $(*)$ есть $o((\|h\| + \|k\|)^2)$. В самом деле,

$$[\Delta_k f]'(a+\theta h) - [\Delta_k f]'(a) = [\Delta_{\theta h}\Delta_k (f')](a).$$

Осталось заметить, что для любой дифференцируемой в точке a функции (в том числе и для $g = f'$) выполняется оценка

$$[\Delta_u \Delta_v g](a) = o(\|u\| + \|v\|)$$

при $u, v \to 0$. (Линейная часть g сокращается, и остаётся $o(\|u\| + \|v\|)$.)

7. Старшие производные

Дифференцируя вторую производную (если это возможно), получаем третью. Значение третьей производной в точке есть элемент $L(X, L(X, L(X, Y)))$; такие операторы можно отождествить с «трилинейными» функциями $X \times X \times X \to Y$ (функциями, линейными по каждому из аргументов при фиксированных остальных).

Из симметрии второй производной легко следует симметрия третьей. В самом деле, вторая производная $f''(x)$ лежит в замкнутом подпространстве симметричных форм, так что и $f'''(a)h$ лежит в этом подпространстве. Другими словами, $[(f'''(a)h)k]l$ симметрично относительно замены $k \leftrightarrow l$. С другой стороны, теорему о симметрии второй производной можно применить к функции f', так что это выражение симметрично относительно замены $h \leftrightarrow k$. Эти две транспозиции порождают все перестановки трёх элементов, так что $f'''(a)$ симметрично.

Аналогичное рассуждение годится и для k-ой производной, которую можно рассматривать как симметричное полилинейное отображение k аргументов

$$X \times X \times \ldots \times X \to Y.$$

8. Формула Тейлора

Формула Тейлора записывается как раньше:

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2!}f''(a)hh + \frac{1}{3!}f'''(a)hhh + \ldots + \\ + \frac{1}{k!}f^{(k)}(a)hh\ldots h + o(\|h\|^k),$$

только теперь вместо h^k мы пишем h подряд k раз, и применяем k-линейную функцию $f^{(k)}(a)$ к набору из k равных векторов. (Или, другими словами: применяем линейный оператор $f^{(k)}(a)$ к вектору h, результат снова применяем к h, результат снова применяем к h и так далее.)

Доказательство также повторяет одномерное. Мы проверяем, что левая и правая части имеют совпадающие производные вплоть до k-ой. Для этого нужно уметь дифференцировать (непрерывную) $x \mapsto p(x) = P(x, x, \ldots, x)$, где P — симметричная полилинейная функция:

$$p'(x)u = kP(u, x, \ldots, x)$$

(где k — число аргументов у функции P). Производная снова есть полилинейная функция со значениями в пространстве линейных (от h) операторов, и вторая производная равна

$$[p''(x)v]u = k(k-1)P(u, v, x, \ldots, x)$$

и так далее.

Проверив, что разность имеет нулевые производные до k-ой включительно, мы затем доказываем по индукции, что $(k-i)$-ая производная разности

есть $o(\|h\|^i)$ (при $i = 1$ по определению, далее с помощью теоремы о конечных приращениях).

Запишем в качестве примера формулу Тейлора для дважды дифференцируемой функции $f\colon \mathbb{R}^n \to \mathbb{R}$ в точке $(0, 0, \ldots, 0)$:

$$f(h_1, \ldots, h_n) = f(0, \ldots, 0) + \sum_i \frac{\partial f}{\partial x_i} h_i + \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} h_i h_j + o(\|h\|^2)$$

Матрицу вторых производных $\frac{\partial f}{\partial x_i \partial x_j}$ называют *гессианом.*

15 Сформулировать формулу Тейлора с оценкой остаточного члена через максимум $(k+1)$-ой производной.

9. Теорема об обратной функции

Если функция f определена и непрерывно дифференцируема в окрестности точки a банахова пространства X и принимает значения в банаховом пространстве Y, причём оператор $f'(a)$ обратим в $L(X, Y)$, то существуют окрестности U точки a и V точки $f(a)$, между которыми f является C^1-диффеоморфизмом. Последнее означает, что f является гомеоморфизмом (непрерывно в обе стороны), притом прямая и обратная функции непрерывно дифференцируемы (*принадлежат классу C^1*, говоря научно).

16 Что это означает при $X = Y = \mathbb{R}$ и почему это верно? (Для $\mathbb{R}^2$ это уже не так очевидно — приходится ссылаться на общее утверждение.)

Доказательство.

(1) *Инъективность*:

$$\|f(x) - f(y) - f'(a)(x - y)\| \leqslant \|f'(\xi) - f'(a)\| \cdot \|y - x\|,$$

а

$$\|f'(a)(x - y)\| \geqslant \|x - y\| / \|f'(a)^{-1}\|,$$

так что поправка к линейной части не может компенсировать линейную часть настолько, чтобы разные точки перешли в одну (если окрестность точки a выбрана достаточно малой).

(2) *Образ содержит окрестность точки $f(a)$*. Пусть b — близкая к $f(a)$ точка пространства Y. Будем искать её прообраз методом последовательных приближений: если текущее приближение z после применения f недобивает до b на $b - f(z)$, то мы вносим поправку, соответствующую линейной части $f'(a)$ (внимание: используется производная в исходной точке a, а не в текущей точке z, как в методе Ньютона):

$$z' = z + f'(a)^{-1}[b - f(z)].$$

Если исходная окрестность точки a достаточно мала, то отображение $z \mapsto z'$ перехода к следующему приближению является сжимающим:

$$z_2' - z_1' = (z_2 - z_1) - f'(a)^{-1}[f(z_2) - f(z_1)],$$

а по теореме о конечном приращении

$$\|f(z_2) - f(z_1) - f'(a)(z_2 - z_1)\| \leqslant (\text{колебание } f') \cdot \|z_2 - z_1\|$$

и достаточно малое колебание f' даже после умножения на $\|f^(a)^{-1}\|$ будет меньше $1/2$, так что отображение будет сжимающим с коэффициентом $1/2$ в любом достаточно малом шаре с центром в a. Если теперь b достаточно близко к $f(a)$, то образ a' точки a удаляется от a не более чем на половину радиуса шара, так что и весь шар переходит в себя и можно использовать теорему о неподвижной точке сжимающего отображения.

(3) *Отображение f открыто.* Предыдущее рассуждение можно применять не только в точке a, но и в любой точке, где производная обратима. Поэтому образ открытого множества открыт, и обратное отображение непрерывно (если ограничить исходное на достаточно малую окрестность точки a).

(4) *Дифференцируемость обратного отображения.* Достаточно проверять её в точке a. Для простоты будем считать, что $X = Y$, $a = f(a) = 0$ и $f'(0) = E$ (общий случай сводится к этому линейной заменой). Условие $f'(0) = E$ означает, что точки x и $y = f(x)$ связаны соотношением

$$y = x + o(\|x\|)$$

при $x \to 0$. А нам нужно проверить, условно говоря, что

$$x = y + o(\|y\|)$$

при $y \to 0$. Уже известная нам непрерывность позволяет заменить $y \to 0$ на $x \to 0$. В малой окрестности $\|o(\|x\|)\| \leqslant \|x\|/2$, и потому $(1/2)\|x\| \leqslant \|y\| \leqslant (3/2)\|x\|$, так что $o(\|y\|)$ и $o(\|x\|)$ — одно и то же.

Зная, что обратное отображение дифференцируемо, мы можем применить теорему о производной композиции и заключить, что производная обратного отображения обратна к $f'(a)$. Как и раньше, это рассуждение можно применить не только в точке a, но и в окрестности a.

(5) *Непрерывность производной обратной функции* следует из непрерывности операции обращения линейного оператора.

Теорема об обратной функции доказана.

10. Теорема о неявной функции

Эта теорема указывает достаточные условия, при которых множество решений уравнения $F(x, y) = 0$ можно локально представить в виде $y = \varphi(x)$ для некоторой непрерывно дифференцируемой функции φ.

Пусть F определена и непрерывно дифференцируема в окрестности точки (a, b) в произведении банаховых пространств $X \times Y$ и принимает значения в банаховом пространстве Z, причём $f(a, b) = 0$, а частная производная f по второй координате Y в точке (a, b) является обратимым линейным оператором (из $L(Y, Z)$). Тогда существуют такие открытые окрестности A у точки a и B у точки b, а также непрерывно дифференцируемая функция $\varphi\colon A \to B$, что в $A \times B$ условия $F(x, y) = 0$ и $y = \varphi(x)$ равносильны.

Пример: для уравнения $F(x, y) = x^2 + y^2 - 1 = 0$ производная $\partial F/\partial y = 2y$ обращается в нуль в двух точках $y = 0, x = \pm 1$. В этих точках теорему о неявной функции применить нельзя, а в окрестности остальных окружность есть график функции $y = \sqrt{1 - x^2}$ (или $y = -\sqrt{1 - x^2}$) в полном соответствии с теоремой.

17 Что даёт теорема о неявной функции для уравнения $y^2 - x^2 = 0$?

Заметим, что окрестность A в утверждении теоремы можно произвольно уменьшать (в отличие от B).

Доказательство. Можно использовать метод доказательства теоремы об обратной функции, решая при каждом x уравнение $F(x, y) = 0$ тем же способом последовательных приближений с использованием $[\partial F/\partial y]^{-1}$. Но можно и формально вывести теорему о неявной функции из теоремы об обратной функции.

Для этого рассмотрим функцию из $X \times Y$ в $X \times Z$, заданную формулой

$$\Phi\colon (x, y) \mapsto (x, F(x, y))$$

(в той же окрестности точки (a, b), что и F). Эта функция непрерывно дифференцируема и производная

$$\Phi'(a, b)\colon (dx, dy) \mapsto \Big(dx, \frac{\partial F}{\partial x}(a, b)\, dx + \frac{\partial F}{\partial y}(a, b)\, dy\Big)$$

обратима одновременно с $\partial F/\partial y$:

$$[\Phi'(a, b)]^{-1}\colon (dx, dz) \mapsto \left(dx, \left[\frac{\partial F}{\partial y}(a, b)\right]^{-1}\Big(dz - \frac{\partial F}{\partial x}(a, b)\, dx\Big)\right)$$

Поэтому существуют окрестности точек (a, b) и $\Phi(a, b) = (a, 0)$, между которыми Φ осуществляет диффеоморфизм. Обозначим их R и S (см. рис. 1).

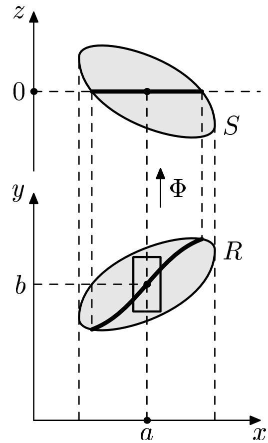


Рис. 1. Обратная ⇒ неявная.

Нас интересует множество решений уравнения $F(x, y) = 0$ внутри R, то есть прообраз подмножества $z = 0$ в $X \times Z$, пересечённого с S, при диффеоморфизме. Теперь видно, что надо положить $\varphi(x)$ равным второй координате $\Phi^{-1}(x, 0)$ (что имеет смысл, если x достаточно близко к a, так как S открыто). Затем надо взять A и B достаточно малыми — настолько, чтобы $A \times B$ было подмножеством R, функция φ была определена в A (то есть $A \times \{0\} \subset S$), после чего надо ещё уменьшить A и добиться $\varphi(A) \subset B$. Последнее можно сделать в силу непрерывности φ, которая гарантируется теоремой об обратной функции. Более того, теорема об обратной функции гарантирует, что φ непрерывно дифференцируема (что нам и требуется). Теорема о неявной функции доказана.

Дифференцируя тождество $F(x, \varphi(x)) = 0$, получаем, что

$$\frac{\partial F}{\partial x}(x, \varphi(x)) + \left[\frac{\partial F}{\partial y}(x, \varphi(x))\right] \varphi'(x) = 0,$$

откуда следует явная формула для производной неявной функции:

$$\varphi'(x) = -\left[\frac{\partial F}{\partial y}(x, \varphi(x))\right]^{-1} \frac{\partial F}{\partial x}(x, \varphi(x)).$$

Производная в квадратной скобке обратима при x, близких к a, поскольку множество обратимых операторов открыто, а производная функции F непрерывна.

Разложение области определения функции F в произведение двух пространств можно выбирать по ходу дела. Пусть дана функция F от n действительных переменных со значениями в $\mathbb{R}^k$, где $k < n$, которая определена и непрерывно дифференцируема в некоторой окрестности точки $a \in \mathbb{R}^n$, причём $F(a) = 0$. Мы интересуемся множеством $F(x) = 0$ (то есть множеством решений системы из k непрерывно дифференцируемых уравнений от n переменных). *Если $F'(a)$ (матрица размера $n \times k$) имеет полный ранг k, то можно так разбить n переменных в $\mathbb{R}^n$ на «независимые» в количестве $n - k$ штук и «зависимые» в количестве k штук, что множество $F(x) = 0$ в окрестности точки a совпадёт с графиком некоторой непрерывно дифференцируемой функции, выражающей зависимые переменные через независимые.*

Легко заметить аналогию с линейной алгеброй, где общий вид решения системы линейных уравнений выражает зависимые переменные через независимые. Доказательство как раз и начинается с этого: мы разбиваем переменные на две группы $\mathbb{R}^{n-k} \times \mathbb{R}^k$, чтобы сужение $F'(a)$ на второе подпространство было обратимо (что соответствует выражению переменных второй группы через первую для линеаризированного уравнения $F'(a)h = 0$). После этого остаётся применить теорему о неявной функции.

11. Подмногообразия в $\mathbb{R}^n$

Говорят, что подмножество $X \subset \mathbb{R}^n$ является *k-мерным C^1-подмногообразием*, если для каждой точки $x \in X$ существует её открытая окрестность U и C^1-диффеоморфизм этой окрестности на открытое множество $V \subset \mathbb{R}^n$, при котором x переходит в 0, а $U \cap X$ переходит в пересечение V с линейным подпространством размерности k.

Примеры: сфера, лента Мёбиуса без края, поверхность конуса без вершины являются двумерными подмногообразиями $\mathbb{R}^3$. (Вершина/край это свойство нарушат, так как у них нет соответствующей окрестности.)

18 Является ли одномерным подмногообразием $\mathbb{R}^2$ кривая, изображённая на рис. 2? (Верхний отрезок не включает в себя крайнюю правую точку; нижняя кривая делает бесконечное число извивов.)

При $k = n$ подмногообразие — это просто открытое множество. При $k = 0$ — это набор изолированных точек.

19 График произвольной непрерывно дифференцируемой функции $f\colon \mathbb{R}^k \to \mathbb{R}^l$ является k-мерным подмногообразием в $\mathbb{R}^{k+l}$.

20 Открытое подмножество подмногообразия является подмногообразием. Всегда ли будет подмногообразием объединение двух непересекающихся подмногообразий?

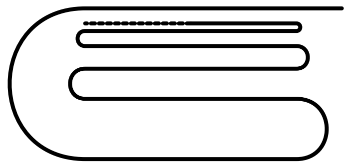


Рис. 2. Подмногообразие?

Если F — непрерывно дифференцируемая функция, определённая в открытом подмножестве $\mathbb{R}^n$ и принимающая значения в $\mathbb{R}^k$ причём $F'(a)$ имеет полный ранг k во всех точках, где $F(x) = 0$, то множество решений системы уравнений $F(x) = 0$ является $(n-k)$-мерным подмногообразием.

(По существу мы уже доказали это в теореме о неявной функции и её следствии.)

Подмногообразия можно задавать не только как прообразы точки при гладких отображениях, но и как образы. Пусть $i\colon U \subset \mathbb{R}^k \to \mathbb{R}^n$ — непрерывно дифференцируемое отображение некоторой открытой окрестности U точки $a \in \mathbb{R}^k$, причём $i'(a)$ как линейное отображение $\mathbb{R}^k \to \mathbb{R}^n$ имеет полный ранг k (является вложением). Тогда *существует открытая окрестность $U' \subset U$, для которой $i(U')$ является k-мерным подмногообразием $\mathbb{R}^n$.*

В самом деле, разобьём координаты в $\mathbb{R}^n$ на две части: k и $n-k$ штук. При этом матрица производной $i'(a)$ разделится на блоки $k \times k$ и $k \times (n-k)$. Разбиение выберем так, чтобы первый из них (минор ранга k) был обратим. (Это возможно, как учит линейная алгебра, поскольку $i'(a)$ имеет ранг k.)

Само отображение i тогда превратится в пару отображений $i_1\colon U \to \mathbb{R}^k$ и $i_2\colon U \to \mathbb{R}^{n-k}$. Первое из них по построению имеет обратимую производную в точке a и по теореме об обратной функции существует окрестность $U' \subset U$, на которой i_1 является диффеоморфизмом. Тогда $f(U')$ будет графиком функции $x_1 \mapsto i_2(i_1^{-1}(x_1))$ и потому подмногообразием.

Замечание. Это утверждение — локальное: образ всего множества U может и не быть многообразием. Например, гладкий путь с не обращающейся в нуль скоростью может самопересекаться.

21 Покажите, что если (в только что описанной ситуации) отображение i является гомеоморфизмом U на $i(U)$, то множество $i(U)$ является подмногообразием. (При этом существенно, что обратное отображение непрерывно; взаимной однозначности недостаточно.)

Если всюду требовать, чтобы отображения были не один раз непрерывно дифференцируемы, а k раз (производные существуют вплоть до k и k-ая производная непрерывна), то получится понятие C^k-подмногообразия.

22 Сформулируйте и докажите аналоги теорем этого раздела для C^k-подмногообразий. [Указание: по существу надо аккуратно повторить те же доказательства.]

12. Касательное пространство

Пусть $X \subset \mathbb{R}^n$ и $x \in X$. Будем рассматривать всевозможные отображения γ окрестностей нуля в $\mathbb{R}$ в пространство $\mathbb{R}^n$, для которых $\gamma(0) = x$ и $\gamma(t) \in X$ при всех t из области определения. (Такие отображения можно назвать путями в X, проходящими через x при $t = 0$.) Среди них выделим дифференцируемые в нуле, и рассмотрим всевозможные векторы $\gamma'(0)$, при этом возникающие. Эти векторы естественно назвать *касательными* к множеству X в точке x.

Если X — подмногообразие, то такие векторы образуют подпространство, размерность которого равна размерности многообразия. В самом деле, при диффеоморфизме касательные векторы переходят в касательные, а для линейного подпространства касательными будут векторы само́го этого подпространства.

Касательное пространство к X в точке x обозначается $T_x(X)$.

23 Покажите, что если X есть график дифференцируемого отображения $f\colon \mathbb{R}^k \to \mathbb{R}^l$, то касательное пространство в точке $(u, f(u))$ этого графика (который представляет собой подмногообразие в $\mathbb{R}^k \times \mathbb{R}^l = \mathbb{R}^{k+l}$) состоит из всех векторов $(h, f'(u)h)$, то есть является графиком производной $f'(u)$ как линейного оператора.

Легко также описать касательное пространство к подмногообразию, заданному как решение системы уравнений или как образ гладкого отображения.

Касательное пространство к множеству решений уравнения $F(x) = 0$ в точке a, представляет собой ядро оператора $F'(a)$. Касательное пространство к образу отображения F из $\mathbb{R}^k$ в $\mathbb{R}^n$ в точке $F(a)$ есть образ оператора $F'(a)$. (В обоих случаях подразумевается, что $F'(a)$ имеет полный ранг, и речь идёт о соответствующем подмногообразии, см. выше)

Для прообраза: если $F(\gamma(t)) = 0$ при всех t, то по формуле дифференцирования сложной функции $\gamma'(0)$ попадает в ядро $F'(a)$. Остаётся заметить, что размерности равны.

Для образа: если рассматривать всевозможные пути в $\mathbb{R}^k$ и их образы, то по теореме о производной сложной функции можно получить любой касательный вектор из образа $F'(a)$. Остаётся снова сослаться на размерности.

13. Карты для подмногообразий

Пусть имеется k-мерное подмногообразие $X \subset \mathbb{R}^n$. *Картой* на X называется взаимно однозначное отображение φ открытого множества $U \subset \mathbb{R}^k$ в X, производная которого (как отображения со значениями в $\mathbb{R}^n$) всюду имеет ранг k.

Если $\varphi(u) = x$, то говорят, что точка x (из подмногообразия) покрыта картой φ и изображается в ней точкой u. Компоненты вектора $u \in \mathbb{R}^k$ называются *координатами* точки x в данной карте.

Метафора понятна: географическую карту можно рассматривать как отображение части бумажной плоскости на часть земной сферы (поверхности земного шара). Географические координаты (широта и долгота) соответствуют карте (в проекции Меркатора, как говорят географы), отображающей прямоугольник $(-180°, 180°) \times (-90°, 90°)$ в сферу; при этом остаётся непокрытой «линия смены дат» и полюса.

Карта покрывает некоторое подмножество подмногообразия X. *Это подмножество открыто в X и карта является гомеоморфизмом.* (Говоря об открытых множествах в X, мы имеем в виду индуцированную из $\mathbb{R}^n$ топологию.)

В самом деле, пусть карта покрывает некоторую точку $x \in X$. По определению подмногообразия существует диффеоморфизм некоторой окрестности точки x (в $\mathbb{R}^n$) на другое открытое подмножество $\mathbb{R}^n$, при котором X переходит в k-мерное подпространство (точнее, в его пересечение с окрестностью). При этом карта превращается в отображение из $\mathbb{R}^k$ в $\mathbb{R}^k$, производная которого имеет ранг k, то есть обратима. Поэтому это отображение открыто, и любая достаточно близкая к x точка покрывается картой. Из этого же следует, что образ открытого множества открыт, так что обратное отображение непрерывно.

Одна и та же область подмногообразия может быть покрыта несколькими картами. Пусть, например, карты $\varphi\colon U \to X$ и $\psi\colon V \to X$ пересекаются:

$$\varphi(U) \cap \psi(V) \neq \varnothing.$$

Тогда возникают взаимно обратные *функции перехода*

$$\varphi^{-1} \circ \psi:\ \psi^{-1}(\varphi(U)) \to \varphi^{-1}(\psi(V))$$

и

$$\psi^{-1} \circ \varphi:\ \varphi^{-1}(\psi(V)) \to \psi^{-1}(\varphi(U)),$$

связывающие изображения одной и той же области $\varphi(U) \cap \psi(V)$ на двух картах (рис. 3).

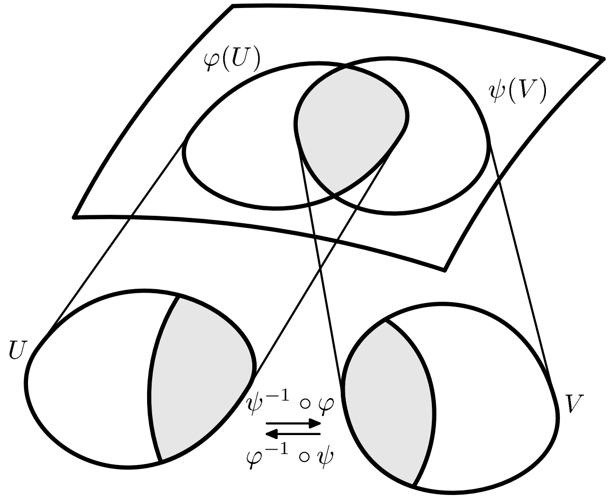


Рис. 3. Функции перехода.

Функции перехода непрерывно дифференцируемы. В самом деле, мы можем локально заменить подмногообразие линейным подпространством размерности k (сделав диффеоморфизм в объемлющем пространстве), после чего сослаться на теорему об обратной функции для k-мерных пространств.

Карты позволяют переходить от k-мерного подмногообразия к открытому множеству в $\mathbb{R}^k$ и обратно. Их можно использовать, чтобы переносить разные понятия анализа на случай многообразий. Скажем, функции f, определённой на подмногообразии X пространства $\mathbb{R}^n$, в карте $\varphi\colon U \to X$ соответствует функция $f\circ\varphi$, определённая уже в U. (При этом часть функции, не попавшая в карту, пропадает — но она может быть видна в другой карте.)

Теперь можно определить понятие дифференцируемой функции, определённой на подмногообразии $X \subset \mathbb{R}^n$ (скажем, на сфере; подчеркнём, что до сих пор такого определения у нас не было). А именно, функция $f\colon X \to \mathbb{R}$ *дифференцируема в точке* $x \in X$, если для любой карты, покрывающей точку x, изображение функции f в карте дифференцируемо в точке, являющейся изображением точки x.

24 Достаточно ли проверить дифференцируемость в одной карте, покрывающей точку x?

25 Докажите, что функция $f\colon X \to \mathbb{R}$ дифференцируема в точке x подмногообразия $X \subset \mathbb{R}^n$ тогда и только тогда, когда она продолжается до

функции, определённой в некоторой $\mathbb{R}^n$-окрестности точки x и дифференцируемой в x.

(Заметим, что мы определили понятие дифференцируемой функции, не определив понятие производной. Естественно ожидать, что в этой ситуации производная будет линейным отображением, но какого пространства? Ответ: $T_x(X)$. См. об этом далее, когда мы будем говорить об абстрактных многообразиях.)

Мы определяли касательный вектор в точке x подмногообразия $X \subset \mathbb{R}^n$ как вектор скорости некоторого пути $t \mapsto \gamma(t)$, проходящего через x при $t = 0$ и не выходящего из X.

В карте $\varphi\colon U \subset \mathbb{R}^k \to X$, покрывающей точку x, пути γ соответствует путь $\varphi^{-1} \circ \gamma$ (возможно, определённый в меньшей окрестности нуля, если γ выходит из области действия карты). Векторы скорости в карте и на подмногообразии связаны отображением $\varphi'(u)$, где u — изображение точки x на карте. Таким образом $\varphi'(u)$ задаёт изоморфизм линейных пространств $\mathbb{R}^k$ и $T_x(X)$; компоненты вектора в $\mathbb{R}^k$ называются *координатами* касательного вектора в данной карте. Поэтому не только точки k-мерного многообразия, но и касательные векторы к нему изображаются в данной карте наборами из k чисел.

14. Теорема о ранге

Пусть U — открытое подмножество $\mathbb{R}^m$ и $f\colon U \to \mathbb{R}^n$ — непрерывно дифференцируемая функция. Тогда для каждой точки $x \in U$ производная $f'(x)$ является линейным оператором $\mathbb{R}^m \to \mathbb{R}^n$ и можно рассмотреть её ранг.

Мы уже рассматривали ситуации, когда $f'(a)$ имеет максимально возможный ранг (равный $\min(m, n)$) в некоторой точке $a \in U$. Для $m < n$ мы доказали, что в этом случае образ некоторой окрестности точки a является подмногообразием в $\mathbb{R}^n$. Для $m > n$ мы доказали, что прообраз точки $f(a)$ является подмногообразием в $\mathbb{R}^m$. Наконец, при $m = n$ отображение f в некоторой окрестности a является диффеоморфизмом (теорема об обратной функции).

Теорема о ранге является обобщением этих утверждений и относится к случаю, когда ранг оператора $f'(a)$ не обязательно максимально возможный. Ему разрешается быть меньше $\min(m, n)$, но мы предполагаем, что в некоторой окрестности точки a ранг $f'(x)$ не больше ранга $f'(a)$ (точка a является «точкой нестрогого локального максимума для ранга»). В этом

случае *существуют*:

$$\begin{array}{ccc} A & \xrightarrow{\ f\ } & B \\ \downarrow & & \downarrow \\ \tilde{A} & \xrightarrow{\ \tilde{f}\ } & \tilde{B} \end{array}$$

- *открытая окрестность A точки a, содержащаяся в U;*
- *открытая окрестность B точки b;*
- *C^1-диффеоморфизм $A \to \tilde{A}$ между A и некоторым открытым множеством $\tilde{A} \subset \mathbb{R}^m$;*
- *C^1-диффеоморфизм $B \to \tilde{B}$ между B и некоторым открытым множеством $\tilde{B} \subset \mathbb{R}^n$,*

для которых отображение $\tilde{f}\colon \tilde{A} \to \tilde{B}$, замыкающее диаграмму, представляет собой линейное отображение ранга k (точнее, его ограничение на множество $\tilde{A}$).

Здесь k — ранг оператора $f'(a)$.

Коротко говоря, в точках локально максимального ранга k все непрерывно дифференцируемые функции устроены (с точностью до диффеоморфизмов) одинаково — как единичная матрица $k \times k$, к которой добавлены нулевые строки и столбцы. В частности, образ малой окрестности точки a, а также прообраз точки $f(a)$ являются (локально) подмногообразиями, что сразу не очевидно.

Набросок доказательства. Пусть $b = f(a)$. В точке a имеется выделенное подпространство Q — ядро отображения $f'(a)$; в точке b выделен образ $f'(a)$, который мы назовём S. Первое подпространство имеет коразмерность k, второе — размерность k. Выделив их прямыми слагаемыми, мы представим оба пространства $\mathbb{R}^m$ и $\mathbb{R}^n$ как произведения $P \times Q = \mathbb{R}^k \times \mathbb{R}^{m-k}$ и $S \times T = \mathbb{R}^k \times \mathbb{R}^{n-k}$ соответственно (см. рис. 4). При этом пространства Q

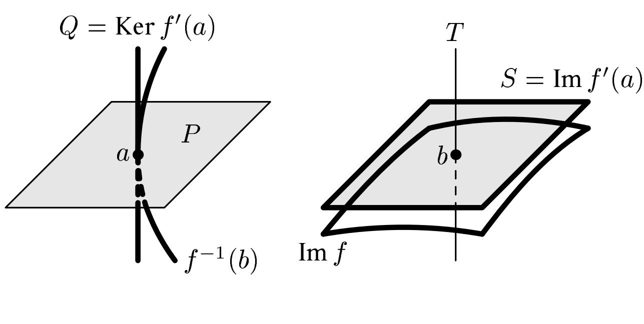


Рис. 4. Разложение в произведение в теореме о ранге.

и S определяются однозначно как ядро и образ оператора $f'(a)$, а в выборе P и T есть произвол.

Будем считать, что точки a и b — пары из нулей в соответствующих пространствах.

Наше отображение теперь можно записать как

$$(p, q) \mapsto (s(p, q), t(p, q))$$

При этом частная производная $\partial s/\partial p$ обратима, и потому отображение $p \mapsto s(p, 0)$ является локальным диффеоморфизмом, а отображение $p \mapsto (s(p, 0), t(p, 0))$ (локально) является картой некоторого k-мерного подмногообразия. В дальнейшем мы увидим, что это подмногообразие совпадает с образом исходного отображения f (образом окрестности точки a в $P \times Q$ при f), но пока это совсем не ясно: непонятно, почему при других (не нулевых) вторых аргументах мы попадём в то же самое подмногообразие.

Аналогичным образом уравнение $s(p, q) = 0$ задаёт подмногообразие в $P \times Q$ (являющееся графиком функции из Q в P по теореме о неявной функции). Впоследствии мы увидим, что это подмногообразие совпадает с прообразом точки $(0, 0)$; уравнение $t(p, q) = 0$ не накладывает новых ограничений, но этого пока мы не знаем.

Решим уравнение

$$s(p, q) = s$$

(мы не вполне корректно используем здесь буквы и как функции, и как точки) относительно p по теореме о неявной функции, выразив (локально) p как функцию от q и s. Другими словами, $p(q, s)$ — это такая точка в P, что

$$s(p(q, s), q) = s,$$

то есть $f(p(q, s), q)$ имеет первую координату s. Мы хотим доказать, что при данном s и различных q все точки $f(p(q, s), s)$ не только проектируются в одну точку s, но даже и совпадают. Для этого надо показать, что производная по q этого выражения равна нулю:

$$\frac{\partial f}{\partial p}(p(q, s), s)\frac{\partial p}{\partial q}(q, s) = 0 \qquad (*)$$

Выражение $\partial f/\partial p$ распадается на две части $\partial s/\partial p$ и $\partial t/\partial p$ в соответствии с разложением пространства в произведение $S \times T$. Теперь смотрите: если оставить только первую часть $\partial s/\partial p$ (заменить f на s в формуле (*)), то получится нуль, поскольку проекция на S равна s и от q не зависит. С другой стороны, в матрице $\partial f/\partial p$ минор $\partial s/\partial p$ имеет максимальный ранг (в силу обратимости в окрестности точки a), и потому если он обращается в нуль на каком-то векторе, то и вся матрица обращается в нуль!

Мы доказали, что (локально) образ функции f является графиком: T-координата определяется S-координатой. Поэтому с помощью диффеоморфизма $(s,t) \mapsto (s, t - t(s))$ его можно сделать плоским. Остаётся в прообразе рассмотреть диффеоморфизм $(p,q) \mapsto (s(p,q), q)$.

(Полноценное рассуждение требует аккуратного выбора окрестностей и т. п., но это мы опускаем.)

15. Абстрактные многообразия

Можно вводить координаты с помощью карт и на произвольном множестве, не обязательно лежащем в координатном пространстве $\mathbb{R}^n$.

Пусть X — некоторое множество. Пусть отображения

$$\varphi_i \colon U_i \to X$$

определены на некоторых открытых множествах $U_i \subset \mathbb{R}^k$ и взаимно однозначно отображают их на некоторые подмножества $\varphi_i(U_i) \subset X$.

Пусть, кроме того, множества $\varphi_i(U_i)$ покрывают всё X, а функции перехода

$$\varphi_j^{-1} \circ \varphi_i \colon \varphi_i^{-1}(\varphi_i(U_i) \cap \varphi_j(U_j)) \to \varphi_j^{-1}(\varphi_i(U_i) \cap \varphi_j(U_j))$$

определены на открытых множествах и непрерывно дифференцируемы. (Меняя местами i и j, получаем обратную функцию перехода, так что функции перехода являются диффеоморфизмами.)

Тогда семейство φ_i называется *k-мерным атласом* на X.

Атлас не обязан быть конечным (индекс i может пробегать произвольное множество). Можно рассматривать *C^m-атласы*, в которых функции перехода являются m раз непрерывно дифференцируемыми; C^1-атласы мы для краткости называем просто атласами.

Два атласа называются *эквивалентными*, если их можно соединить в один, то есть функции перехода из карт первого в карты второго и наоборот определены на открытых множествах и непрерывно дифференцируемы. (Предупреждение: на одном и том же абстрактном множестве мощности континуум можно задать самые разные атласы, превращающие его в многообразия разных размерностей.)

Многообразием размерности k называется множество X, снабжённое k-мерным атласом, при этом эквивалентные атласы не различаются. (Формально: многообразие есть множество с классом эквивалентных атласов.)

26 Карта согласована с атласом, если её добавление даёт новый атлас (то есть функции перехода между новой картой и любой старой хорошие).

Покажите, что если добавить к атласу все согласованные с ним карты, то получится максимальный (по включению) из эквивалентных атласов.

На многообразии естественно возникает структура топологического пространства. А именно, множество $V \subset X$ считается открытым, если открыто его изображение (точнее, изображение его части) в любой карте, то есть $\varphi^{-1}(V)$ открыто в $\mathbb{R}^k$ для любой карты φ нашего атласа.

27 Проверьте, что это определение корректно, то есть не меняется при замене атласа на эквивалентный.

28 Проверьте, что для любой карты $\varphi\colon U \to X$ образ любого открытого множества $V \subset U$ будет открытым подмножеством X. (В частности, $\varphi(U)$, то есть «зона действия» карты φ, является открытым в X множеством.) Проверьте, что каждая карта $\varphi : U \to \varphi(U)$ является гомеоморфизмом (с точки зрения индуцированных топологий в U из $\mathbb{R}^k$ и в $\varphi(U)$ из X).

С замкнутыми множествами сложнее: если множество целиком содержится в одной карте и изображается в ней замкнутым множеством, то оно не обязательно замкнуто (например, пусть карта отображает целиком $\mathbb{R}^k$ в X и множество есть образ этой карты). Но множество, содержащееся в одной карте и изображающееся в ней компактом, компактно (поскольку карта — гомеоморфизм).

Предупреждение. Данное определение не исключает патологические случаи. Например, можно взять две копии прямой и отождествить отрицательные числа в обеих копиях (так что отрицательных чисел имеется по одному, а нуль и положительные — в двух экземплярах). При этом возникают две карты, функции перехода — тождественные на открытом множестве отрицательных чисел, и все требования выполнены.

Чтобы избежать этого, мы будем всегда требовать, чтобы многообразие было отделимым топологическим пространством (две различные точки имеют непересекающиеся окрестности). В нашем примере многообразие неотделимо (две копии нуля не имеют непересекающихся окрестностей; указанное свойство заведомо выполнено для двух точек, если они лежат в одной карте или лежат в непересекающихся картах, но с двумя копиями нуля это не так).

Примеры многообразий.

1. Подмногообразия $\mathbb{R}^n$. (Понятие карты для подмногообразия мы уже обсуждали.)

2. Множество всех прямых на плоскости. (Можно взять две карты — одна покрывает негоризонтальные прямые, а другая — невертикальные.) Множество всех прямых или всех плоскостей в пространстве. Множество k-мерных подпространств в векторном (или аффинном) n-мерном пространстве.

3. *Конфигурационное пространство* механической системы, скажем, множество положений маятника (окружность), пары маятников (тор), кругового маятника (сфера), тела в пространстве ($\mathbb{R}^3 \times SO(3)$) или что-нибудь в таком роде. (Важно только, чтобы не было положений, когда одна часть системы упирается в другую — тогда в их окрестности не получится карты.)

4. Множество всех касательных прямых к данной сфере.

В каждом из этих примеров аккуратное определение требует описания карт и проверки гладкости функций перехода (а также отделимости).

Подробнее введённые нами многообразия (с непрерывно дифференцируемыми функциями перехода) называются C^1*-многообразиями*. Аналогично можно рассматривать C^m*-многообразия* (требуя, чтобы функции перехода были m раз непрерывно дифференцируемыми), C^∞*-многообразия* (с бесконечно дифференцируемыми функциями перехода), называемые также *гладкими многообразиями*, *аналитические* многообразия (с аналитическими функциями перехода) и др. Можно также говорить о *комплексных* многообразиях (предварительно определив дифференцирование для функций с комплексными аргументами и значениями).

29 Укажите структуру многообразия на произвольном открытом подмножестве многообразия.

30 Укажите размерности многообразий в приведённых выше примерах. Предложите какие-нибудь «естественные» карты для них и проверьте гладкость функций перехода.

31 Многообразие называется *ориентируемым*, если существует атлас, в котором все функции перехода имеют положительный определитель производной (*якобиан*) во всех точках. Покажите, что сфера ориентируема, а лист Мёбиуса — нет.

16. Гладкие функции. Касательные векторы. Производная

Пусть имеются два многообразия X и Y и функция $f\colon X \to Y$. Говорят, что f *дифференцируема* в точке $x \in X$, если для любых карт $\varphi\colon U \to X$ и $\psi\colon V \to Y$, покрывающих точки x и $f(x)$ соответственно, изображение функции f в этих картах, то есть функция

$$\psi^{-1} \circ f \circ \varphi\colon \varphi^{-1}(f^{-1}(\psi(V))) \to V$$

определена в окрестности точки $\varphi^{-1}(x)$ и дифференцируема в этой точке.

32 Покажите, что определение дифференцируемости не зависит от выбора карты и достаточно проверить его для каких-нибудь карт, покрывающих x и $f(x)$.

33 Покажите, что из дифференцируемости функции в точке многообразия следует её непрерывность (в описанной выше топологии на многообразии).

34 Дайте определение непрерывно дифференцируемой функции. Дайте определение k раз непрерывно дифференцируемой функции для C^m-многообразий при $m \geqslant k$.

35 Покажите, что для функций, определённых на открытых множествах в $\mathbb{R}^n$ и принимающих значения в подмногообразиях $\mathbb{R}^m$, новое определение совпадает со старым (для функций из $\mathbb{R}^n$ в $\mathbb{R}^m$).

36 (**а**) Докажите, что если точка $x \in \mathbb{R}^n$ не принадлежит замкнутому множеству $X \subset \mathbb{R}^n$, то существует гладкая функция, равна нулю в окрестности точки x и равная единице на X.

(**б**) Докажите аналогичное утверждение для произвольного гладкого многообразия M (вместо $\mathbb{R}^n$).

(**в**) Докажите, что если на многообразии M имеются два непересекающихся замкнутых множества A и B, причём A компактно, то можно построить гладкую функцию, которая всюду положительна на A и равна нулю на B. Как из неё получить гладкую функцию, равную 1 на A и 0 на B, все значения которой принадлежат отрезку $[0, 1]$?

(**г**) (разбиение единицы) Докажите, что для любого набора открытых множеств $U_1, \ldots, U_k$, покрывающих компактное множество C на многообразии M, найдутся такие гладкие функции $f_1, \ldots, f_k \colon M \to [0, 1]$, что что (1) носитель функции f_i (замыкание множества тех точек, где $f_i \neq 0$) компактен и лежит в U_i; (2) $f_1(x) + \ldots + f_k(x) = 1$ при всех $x \in C$.

Заметим, что пока что у нас есть определение дифференцируемой функции, но неизвестно, что считается её производной. Чтобы определить производную, сначала определим касательное пространство (к многообразию в его точке).

Пусть X — некоторое многообразие, а x — его точка. Будем рассматривать всевозможные пути, проведённые через x, то есть дифференцируемые отображения

$$\gamma \colon (-\varepsilon, \varepsilon) \to X$$

(при всевозможных $\varepsilon > 0$), для которых $\gamma(0) = x$. Два таких пути γ_1, γ_2 будем называть эквивалентными («имеющими одинаковое направление в точке x»), если их изображения в некоторой карте φ, то есть пути $\varphi^{-1} \circ \gamma_1$ и

$\varphi^{-1} \circ \gamma_2$, определённые в некоторой (возможно, меньшей) окрестности нуля, имеют одинаковую производную в нуле.

37 Проверьте, что если условие эквивалентности путей верно для некоторой карты, то оно верно и для любой другой согласованной карты.

Класс эквивалентных путей называется *касательным вектором в точке x к многообразию X*.

Производная в точке 0 пути $\varphi^{-1} \circ \gamma$ по определению не зависит от выбора γ в классе эквивалентности; этот вектор называется *изображением* касательного вектора в карте φ. Легко понять, что изображение может быть любым вектором (возьмём путь в карте и перенесём его в многообразие). Тем самым имеется взаимно однозначное соответствие между изображениями (элементами $\mathbb{R}^k$) и касательными векторами, и на векторах возникает структура векторного пространства, называемого *касательным пространством в точке x к многообразию X*.

38 Почему эта структура векторного пространства не зависит от выбора карты?

Пусть имеются две карты, покрывающие одну и ту же точку x на многообразии. Как связаны изображения одного и того же вектора в этих картах? Соответствующие пути связаны функцией перехода, поэтому векторы скорости связаны её производной.

Поэтому можно сказать, что касательный вектор в точке x — это когда в каждой карте задан некоторый вектор в $\mathbb{R}^k$, причём векторы для разных карт связаны производной функции перехода в изображении точки x. Или на ещё более «физическом» языке: касательный вектор — это набор чисел, умножающийся при переходе к другой системе координат на матрицу Якоби частных производных функций преобразования координат из первой системы во вторую.

Пусть f — дифференцируемая функция, определённая в окрестности точки x в многообразии X и принимающая значения в некотором нормированном пространстве Y. Пусть v — касательный вектор к X в точке x. Определим *производную функции f вдоль вектора v* (обозначение: $D_v f$) как производную функции

$$t \mapsto f(\gamma(t))$$

в точке 0 для некоторого представителя γ вектора v.

39 Проверьте, что это определение корректно, то есть что производная не зависит от выбора представителя.

Отображение

$$v \mapsto D_v f$$

является линейным отображением касательного пространства $T_x(X)$ в Y. В самом деле, структура линейного пространства определяется с помощью карты, а после перехода в карту φ получается отображение

$$w \mapsto \text{производная } f\circ\varphi \text{ вдоль } w$$

(производная функции f вдоль вектора v равна производной функции $f\circ\varphi$ вдоль изображения w вектора v в карте φ).

Построенное отображение $T_x(X) \to Y$ называется *производной* (или *дифференциалом*) функции f в точке x и обозначается $f'(x)$.

40 Пусть f и g — функции с действительными значениями, определённые в окрестности точки x многообразия X и дифференцируемые в точке x. Докажите *правило Лейбница*:

$$D_v(fg) = f(x)(D_v g) + (D_v f)g(x).$$

На этом наблюдении основано алгебраическое определение касательных векторов.

Рассмотрим множество гладких (бесконечно дифференцируемых) функций, определённых в окрестностях точки $0\in\mathbb{R}^n$. Будем считать две такие функции эквивалентными, если они совпадают в некоторой окрестности точки 0 (эта окрестность, естественно, должна содержаться в их областях определения). Класс эквивалентных функций назовём *ростком* (гладкой функции в нуле).

41 Определите структуру линейного пространства и кольца на множестве ростков гладких функций в нуле.

Пусть $\mathscr{S}$ — кольцо ростков гладких функций в нуле (в $\mathbb{R}^n$). *Дифференцированием* называется линейное отображение $D\colon \mathscr{S}\to\mathbb{R}$, для которого

$$D(fg) = f(0)D(g) + D(f)g(0).$$

Предыдущая задача показывает, что для любого вектора $v\in\mathbb{R}^n$ отображение

$$D_v(f) = f'(0)v$$

(производная в нуле вдоль вектора v) является дифференцированием. Верно и обратное утверждение: других дифференцирований (кроме D_v) нет. В самом деле, сначала найдём производную констант:

42 Докажите, что $D(c) = 0$ для любого дифференцирования D и любой константы (постоянной функции) c. (Указание. По линейности достаточно проверить $D(1) = 0$, что получается из формулы Лейбница для $D(1\cdot 1)$.)

Затем нужно воспользоваться *леммой Адамара*: всякая гладкая функция f, определённая в окрестности нуля в $\mathbb{R}^n$, представима в окрестности нуля как

$$f(x_1,\ldots,x_n)=f(0,\ldots,0)+x_1g_1(x_1,\ldots,x_n)+\ldots+x_ng_n(x_1,\ldots,x_n)$$

для некоторых гладких функций $g_1,\ldots,g_n$.

В одномерном случае лемма Адамара по существу утверждает, что для гладкой функции f в окрестности нуля частное $(f(x)-f(0))/x$, доопределённое в нуле по непрерывности как $f'(0)$, является гладкой функцией. (Попробуйте это доказать непосредственно.) В многомерном случае разложение, даваемое леммой Адамара, уже не однозначно. Для аналитических функций (в частности, многочленов) надо сгруппировать слагаемые с x_1 в одну часть, с x_2 в другую и так далее (при этом слагаемое x_1x_2 может быть отнесено куда угодно). Общее доказательство леммы Адамара использует интеграл и возможность дифференцирования под знаком интеграла: рассмотрим для фиксированного $x\in\mathbb{R}^n$ функцию $\alpha(t)=f(tx)$ при $t\in[0,1]$. Тогда

$$f(x)=f(0)+\alpha(1)-\alpha(0)=f(0)+\int_0^1\alpha'(t)\,dt$$

Поскольку $\alpha'(t)=f'(tx)x=(\operatorname{grad}f(tx),x)$, то интеграл можно переписать как сумму (по $i=1,\ldots,n$) слагаемых

$$x_i\int_0^1\frac{\partial f}{\partial x_i}(tx_1,\ldots,tx_n)\,dt=x_ig_i(x_1,\ldots,x_n)$$

что и даст искомое представление для f, если поверить в гладкость функций g_i (она вытекает из общих теорем о дифференцировании под знаком интеграла).

Применяя дифференцирование D к обеим частям леммы Адамара, находим, что

$$Df=g_1(0)Dx_1+\ldots+g_n(0)Dx_n,$$

где Dx_i есть число, получаемое применением D к координатной функции $(x_1,\ldots,x_n)\mapsto x_i$. Тот же результат даёт дифференцирование вдоль вектора $(v_1,\ldots,v_n)$, если положить $v_i=Dx_i$.

43 Убедитесь, что аналогичные рассуждения применимы к дифференцированию ростков гладких функций на многообразии класса C^∞ и тем самым касательный вектор к такому многообразию можно определять как дифференцирование кольца ростков.

Вернёмся к обычному определению касательного вектора как класса эквивалентных путей и объясним, что такое производная дифференцируемого

отображения одного многообразия в другое. Пусть X и Y — два многообразия и f — отображение открытого множества $U \subset X$ в Y, дифференцируемое в точке $x \in U$.

Определим производную отображения f, которая будет линейным отображением касательных пространств $T_x(X) \to T_{f(x)}Y$. А именно, если $v \in T_x(X)$ есть класс некоторого пути γ, рассмотрим путь $f \circ \gamma$, проходящий через $f(x)$, и класс этого пути.

44 Проверьте, что это определение корректно. Проверьте, что в картах это определение совпадает с обычным определением производной. Сформулируйте и докажите теорему о производной композиции двух отображений многообразий.

Таким образом, в определении производной отображения f одного многообразия в другое уже не осталось дифференцирования как такового, а есть лишь композиция путей с f; дифференцирование скрыто в определении векторов как классов эквивалентных путей (эквивалентность определяется как равенство производных).

Более «физическое» определение производной как отображения касательных пространств выглядело бы так: матрица частных производных изображения функции $f: X \to Y$ в картах зависит от выбора карт на X и на Y. Если в X имеется набор чисел, зависящих от выбора карты как положено для вектора, то при умножении этого вектора на упомянутую матрицу зависимость от карты в X пропадёт, и результат является набором чисел, зависящим от выбора Y-карты надлежащим образом, то есть образует касательный вектор к Y.

17. Экстремумы на многообразиях

Пусть функция f, определённая на многообразии X, принимает действительные значения. *Если функция f достигает максимума в точке $a \in X$ и дифференцируема в a, то $f'(a) = 0$.*

На всякий случай поясним: «достигает максимума» означает, что $f(a) \geqslant f(x)$ при всех $x \in X$. В равенстве $f'(a) = 0$ нуль в правой части есть нулевое линейное отображение пространства $T_a(X)$ в $\mathbb{R}$.

В самом деле, пусть путь $\gamma: (-\varepsilon, \varepsilon) \to X$ представляет касательный вектор в точке a; в частности, тогда $\gamma(0) = a$. По определению, образ этого вектора при $f'(a)$ есть производная функции $t \mapsto f(\gamma(t))$ в нуле, которая равна нулю, поскольку здесь речь идёт о действительной функции действительного аргумента, имеющей максимум в нуле.

Очевидно, то же самое верно и для *локального* максимума или минимума.

Частный случай: пусть X является k-мерным подмногообразием в $\mathbb{R}^n$, а функция f определена и дифференцируема в окрестности подмногообразия X, то есть на некотором открытом в $\mathbb{R}^n$ множестве, содержащем всё X. Задача о поиске максимума (минимума) на X тогда называется задачей об *условном экстремуме*, поскольку нас интересует максимум или минимум только при условии попадания аргумента в X. Соответственно и производная $f'(a)$ как линейный оператор $\mathbb{R}^n \to \mathbb{R}$ обязана обращаться в нуль не целиком, а лишь на k-мерном подпространстве $T_a(X) \subset \mathbb{R}^n$.

Та же ситуация ещё более конкретно: пусть функция

$$f\colon (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$$

достигает максимума в точке a на подмногообразии, заданном уравнениями

$$\begin{aligned} g_1(x_1, \ldots, x_n) &= 0 \\ &\ldots \\ g_s(x_1, \ldots, x_n) &= 0, \end{aligned}$$

причём матрица $\|\partial g_i / \partial x_j\|$ имеет максимальный ранг s (так что размерность подмногообразия равна $n - s$). Тогда сформулированное выше условие превращается в такое:

$$g_1'(a)h = 0, \ldots, g_s'(a)h = 0 \ \Rightarrow f'(a)h = 0$$

для любого вектора $h \in \mathbb{R}^n$.

Лемма из линейной алгебры: если $\alpha_1, \ldots, \alpha_s, \beta$ — функционалы на векторном пространстве E, и при этом

$$\ker \alpha_1 \cap \ldots \cap \ker \alpha_s \subset \ker \beta,$$

то β линейно выражается через $\alpha_1, \ldots, \alpha_s$, то есть существуют такие числа $\lambda_1, \ldots, \lambda_s$, что

$$\beta = \lambda_1 \alpha_1 + \ldots + \lambda_s \alpha_s.$$

(Доказательство: предположение гарантирует, что $\beta(e)$ однозначно определяется значениями $\alpha_i(e)$; остаётся продолжить эту линейную зависимость с образа оператора $(\alpha_1, \ldots, \alpha_s)$ на всё $\mathbb{R}^s$.)

Таким образом, если a является точкой условного максимума функции f при условиях $g_1 = \ldots = g_s = 0$, то существуют такие числа $\lambda_1, \ldots, \lambda_s$, что

$$f'(a) - \lambda_1 g_1'(a) - \ldots - \lambda_s g_s'(a) = 0,$$

причём это равенство выполнено на всём $\mathbb{R}^n$. Числа λ_i называются *множителями Лагранжа*.

Другими словами: если мы ищем максимум выражения $f(x_1, \ldots, x_n)$ без ограничений на $(x_1, \ldots, x_n)$, то надо записать n уравнений $\partial f/\partial x_i(x_1, \ldots, x_n) = 0$ (при $i = 1, \ldots, n$). Если появляется s ограничений $g_i(x_1, \ldots, x_n) = 0$ (при $i = 1, \ldots, s$), то надо добавить неопределённые коэффициенты $\lambda_1, \ldots, \lambda_s$ и записать те же n уравнений (обращение в нуль частных производных) для функции

$$f(x_1, \ldots, x_n) - \lambda_1 g_1(x_1, \ldots, x_n) - \ldots - \lambda_s g_s(x_1, \ldots, x_s).$$

Получится n уравнений с $n+s$ неизвестными $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_s$, к которым добавляются ещё s уравнений $g_i(x_1, \ldots, x_n) = 0$.

Пример 1. Докажем неравенство о среднем арифметическом и среднем геометрическом в таком варианте: если $x_1, \ldots, x_n \geqslant 0$ и $x_1 + \ldots + x_n = n$, то $x_1 \cdot x_2 \cdot \ldots \cdot x_n \leqslant 1$.

В самом деле, функция $f(x_1, \ldots, x_n) = x_1 \cdot x_2 \cdot \ldots \cdot x_n$ непрерывна и потому достигает максимума на компактном множестве S, заданном условиями $x_1, \ldots, x_n \geqslant 0$ и $x_1 + \ldots + x_n = n$. Этот максимум больше нуля и потому достигается во внутренней точке (на границе функция обращается в нуль). Внутренние точки образуют подмногообразие, заданное уравнением $g = 0$, где $g(x_1, \ldots, x_n) = x_1 + \ldots + x_n - n$. По правилу множителей Лагранжа запишем

$$\frac{\partial}{\partial x_i}[f(x_1, \ldots, x_n) - \lambda g(x_1, \ldots, x_n)] = 0,$$

что в нашем случае означает

$$x_1 \cdot \ldots \widehat{x_i} \ldots \cdot x_n - \lambda = 0$$

(символ $\widehat{x_i}$ означает, что сомножитель x_i пропущен). Во внутренней точке все x_i положительны, поэтому отсюда следует, что они равны и максимум достигается в точке $(1, \ldots, 1)$. Что и требовалось.

Пример 2. Докажем, что любая симметрическая действительная матрица A размера $n \times n$ имеет собственный вектор. (Далее можно доказать, что она имеет и собственный базис из ортогональных векторов, но этого мы касаться не будем.)

Для доказательства рассмотрим максимум выражения

$$f(x_1, \ldots, x_n) = (Ax, x) = \sum_{i,j} a_{ij} x_i x_j$$

(сумма берётся по всем n^2 парам (i, j)) на единичной сфере

$$g(x_1, \ldots, x_n) = x_1^2 + \ldots + x_n^2 - 1 = 0.$$

Поскольку функция непрерывна, а сфера компактна, то максимум достигается в некоторой точке сферы. В этой точке для некоторого λ все частные производные функции $f - \lambda g$ обращаются в нуль, что даёт

$$\frac{\partial}{\partial x_k}[f - \lambda g] = 2\sum_i a_{ki}x_i - 2\lambda x_k = 0$$

(при дифференцировании нужно отдельно учесть диагональные члены $a_{ii}x_i^2$, а также пары симметричных внедиагональных членов $a_{ik}x_ix_k$ и $a_{ki}x_kx_i$, в обоих случаях появляется множитель 2). А это в точности означает, что вектор $(x_1, \ldots, x_n)$ является собственным вектором оператора с матрицей A с собственным значением λ.

45 Проведите это рассуждение без координат, дифференцируя квадратичные формы (Ax, x) и (x, x).

46 Докажите, что диагональ прямоугольного параллелепипеда, описанного в евклидовом пространстве вокруг эллипсоида, не зависит от выбора параллелепипеда. [Указание. Удобно считать, что стороны параллелепипеда параллельны осям координат, а эллипсоид $(Ax, x) = 1$ вращается. Вращение эллипсоида не меняет следа матрицы A^{-1}, через который выражается диагональ параллелепипеда.]

18. Дифференциальные формы первой степени

Пусть функция $f\colon X \to \mathbb{R}$ дифференцируема во всех точках многообразия X. Тогда для каждой точки $x \in X$ производная $f'(x)$ является линейным функционалом на касательном пространстве. Такие отображения — когда каждой точке $x \in X$ ставится в соответствие линейный функционал $\omega(x)\colon T_x(X) \to \mathbb{R}$ — называются *дифференциальными формами первой степени* на многообразии X. Таким образом, каждой дифференцируемой функции f соответствует дифференциальная форма, которая называется её *дифференциалом* и обозначается df.

Если X представляет собой открытое множество в $\mathbb{R}^n$, то для каждой точки $x = (x_1, \ldots, x_n) \in X$ значение дифференциальной формы ω в этой точке есть линейный функционал $\omega(x)\colon \mathbb{R}^n \to \mathbb{R}$ и задаётся своими координатами в стандартном базисе:

$$\omega(x) = \omega_1(x)e^1 + \ldots + \omega_n(x)e^n.$$

Здесь e^i — функционал проекции на i-ю координату. Поскольку этот функционал является производной (линейной) функции $(x_1, \ldots, x_n) \mapsto x_i$ в лю-

бой точке, то его обозначают dx_i и разложение формы ω по базису записывают в виде

$$\omega(x) = \omega_1(x)\,dx_1 + \ldots + \omega_n(x)\,dx_n.$$

Форма ω называется *непрерывной*, если непрерывны функции ω_i. Это определение непрерывности годится для форм на открытых подмножествах $\mathbb{R}^n$; для форм на произвольном многообразии надо сначала определить (естественным образом) изображение формы в карте, а затем сказать, что форма непрерывна, если её изображения во всех картах непрерывны.

47 Сформулируйте это аккуратно и проверьте, что непрерывность дифференциальной формы первой степени не меняется при замене атласа на C^1-эквивалентный.

В дальнейшем мы рассматриваем только непрерывные дифференциальные формы, не оговаривая этого особо.

Форма первой степени ω на многообразии X называется *точной*, или *кограницей*, если она является дифференциалом некоторой функции f (которая тем самым будет непрерывно дифференцируемой): $\omega = df$. Эта функция f называется *первообразной* для формы ω.

48 Рассмотрим на множестве $x \neq 0$ (в $\mathbb{R}^2$) функцию $\varphi(x,y) = \operatorname{arctg}(y/x)$. Найдите её дифференциал и покажите, что он продолжается до непрерывной формы на $\mathbb{R}^2 \setminus \{(0,0)\}$.

Для случая интервала на прямой дифференциальная форма имеет вид $P(x)\,dx$; её первообразная находится интегрированием $\int P(x)\,dx$ и существует для любой непрерывной формы. В общем случае это уже не так, как мы увидим, и препятствия к существованию первообразной могут быть разными (локальными и глобальными). Но и здесь начнём с определения интеграла.

19. Криволинейный интеграл и кограницы

Пусть форма ω определена на многообразии X и имеется непрерывно дифференцируемый путь $\gamma\colon [a,b] \to X$ (это означает, что γ продолжается до непрерывно дифференцируемого отображения некоторого интервала, содержащего отрезок $[a,b]$). Определим *криволинейный интеграл формы ω по пути γ* формулой

$$\int_\gamma \omega = \int_a^b \omega(\gamma(t))\gamma'(t)\,dt.$$

Здесь $\omega(\gamma(t))$ есть значение дифференциальной формы ω в точке $\gamma(t)$, то есть линейный функционал на касательном пространстве в точке $\gamma(t)$, а

$\gamma'(t)$ — вектор того самого пространства, так что выражение под интегралом есть вещественная функция от t. Эта функция непрерывна, так как форма непрерывна по предположению, а путь непрерывно дифференцируем. Поэтому интеграл (Римана) имеет смысл.

Основные свойства криволинейного интеграла:

• если путь γ разбит на две части γ_1 и γ_2 точкой $c \in [a, b]$, то интеграл по γ равен сумме интегралов по γ_1 и γ_2;

• криволинейный интеграл не меняется при изменении параметризации пути: если отображение φ представляет собой диффеоморфизм окрестности отрезка $[a, b]$ на окрестность отрезка $[\tilde{a}, \tilde{b}]$, при котором a и b переходят в $\tilde{a}$ и $\tilde{b}$ соответственно, а пути $\tilde{\gamma}\colon [\tilde{a}, \tilde{b}] \to X$ и $\gamma\colon [a, b] \to X$ связаны равенством $\gamma = \tilde{\gamma} \circ \varphi$, то интегралы от любой формы по путям γ и $\tilde{\gamma}$ одинаковы;

• при изменении направления пути (см. предыдущее утверждение, но $\varphi(a) = \tilde{b}$ и $\varphi(b) = \tilde{a}$) интеграл по пути меняет знак;

• интеграл от дифференциала df непрерывно дифференцируемой функции f по пути γ равен разности значений функции f в конце и начале пути (*формула Ньютона – Лейбница*).

Все эти свойства следуют из аналогичных свойств интеграла Римана. Например, второе и третье соответствуют правилу замены переменной в определённом интеграле, а последнее — формуле Ньютона – Лейбница: ведь $df(\gamma(t))\gamma'(t)$ есть как раз производная функции $t \mapsto f(\gamma(t))$ по теореме о производной сложной функции.

Первое из перечисленных свойств позволяет определить интеграл по кусочно-гладкому пути. Непрерывный путь $[a, b] \to X$ на многообразии X называется *кусочно-гладким*, если отрезок $[a, b]$ можно разбить на конечное число меньших отрезков, на каждом из которых путь является непрерывно дифференцируемым (то есть продолжается до непрерывно дифференцируемого отображения большего интервала). Тогда интеграл по всему пути определяется как сумма интегралов по его частям, и первое из перечисленных свойств позволяет доказать корректность этого определения (если есть два разбиения, рассмотрим их общее измельчение).

В частности, можно рассматривать интегралы по границам прямоугольников, это нам пригодится.

49 Убедитесь, что формула Ньютона – Лейбница верна и для кусочно-гладких путей.

Критерий точности формы в терминах интеграла. *Следующие три свойства дифференциальной формы ω первой степени на многообразии X равносильны:*

• *интегралы формы ω по двум кусочно-гладким путям с общими началом*

и концом равны;

• *интеграл формы ω по любой кусочно-гладкой петле, то есть пути с совпадающим началом и концом, равен нулю*;

• *форма ω точна.*

Первое и второе свойства равносильны. Если интеграл зависит только от начала и конца, то интеграл по петле равен интегралу по постоянному пути, то есть нулю. Если интеграл по любой петле равен нулю, то соорудим петлю из двух путей с общим началом и концом, изменив направление в одном из них. (Здесь важно, что разрешаются кусочно-гладкие пути.)

Из третьего свойства следуют два первых по формуле Ньютона – Лейбница. Остаётся доказать, что если интеграл формы зависит лишь от начала и конца пути, то форма точна. Фиксируем произвольно точку a на многообразии и определим $f(x)$ как интеграл от ω по любому пути, ведущему из a в x. (Мы предполагаем, что многообразие X линейно связно; если это не так, можно рассматривать его компоненты по отдельности.) Независимость интеграла от пути гарантирует, что это определение корректно. (Заметим также, что выбор точки a меняет функцию f лишь на константу.)

Остаётся проверить, что $df = \omega$. Это утверждение локально, поэтому его можно проверять в карте (и при этом можно считать, что начальная точка попадает в ту же карту). Пусть имеется форма с непрерывными коэффициентами

$$P_1(x_1, \ldots, x_n)\, dx_1 + \ldots + P_n(x_1, \ldots, x_n)\, dx_n$$

Достаточно проверить, что частные производные функции f, полученной интегрированием этой формы, равны P_i (поскольку форма непрерывна по предположению, то P_i непрерывны и отсюда будет следовать существование полной производной функции f).

Разность значений $f(x_1, \ldots, x_i + \Delta x_i, \ldots, x_n)$ и $f(x_1, \ldots, x_i, \ldots, x_n)$ равна интегралу по отрезку, идущему вдоль x_i, то есть

$$\int_{x_i}^{x_i + \Delta x_i} P_i(x_1, \ldots, x_{i-1}, t_i, x_{i+1}, \ldots, x_n)\, dt_i$$

поэтому достаточно сослаться на одномерную формулу Ньютона – Лейбница.

50 Проверьте, что формы $x\,dx + 2y\,dy$ и $y\,dx + x\,dy$ являются точными формами (и укажите первообразную), а $2y\,dx + x\,dy$ — нет.

20. Замкнутые формы (коциклы)

Что мешает дифференциальной форме быть точной? Возможны препятствия двух сортов: локальные (форма не является точной даже в малой

окрестности) и глобальные (в малых окрестностях можно найти первообразную, но они не склеиваются). Будем изучать эти вопросы раздельно.

Дифференциальная форма ω на многообразии X является *замкнутой*, или *коциклом*, если локально она точна: для каждой точки $a \in X$ существует окрестность U точки a и такая непрерывно дифференцируемая функция $f\colon U \to \mathbb{R}$, что $df = \omega$ в U.

Не всякая замкнутая форма точна. Стандартный пример замкнутой, но не точной формы: $d\varphi$, определяемая как дифференциал функции полярного угла. Полярный угол, строго говоря, не есть функция — он определён лишь с точностью до 2π (во всех точках плоскости, кроме начала координат). Формально его можно определить как $\operatorname{arctg}(y/x) + 2\pi k$ (при $x \neq 0$) или как $\pi/2 - \operatorname{arctg}(x/y) + 2\pi k$ (при $y \neq 0$). Поскольку прибавление 2π не меняет производной, то производная определена корректно на плоскости без точки (см. задачу 48, где указано выражение для $d\varphi$).

Из этого определения видно, что форма $d\varphi$ точна в областях $x \neq 0$ и $y \neq 0$, и потому замкнута. Однако её интеграл по единичному кругу равен 2π и потому она не точна.

51 Вычислите этот интеграл для пути $(\cos t, \sin t)$ явно, а также с помощью формулы Ньютона–Лейбница. (Указание: в последнем случае удобно разбить путь на части.)

Прежде чем указать критерий замкнутости для гладких форм, сформулируем лемму, которая пригодится в его доказательстве.

Лемма. Если форма определена в круге и интеграл по любому прямоугольнику внутри этого круга равен нулю, то форма точна.

В самом деле, из условия леммы следует, что интегралы по двум путям, ведущим из центра круга в точку x (см. рис. 5), одинаковы. Определим первообразную функцию f как интеграл по любому из них. Путь вправо–вверх удобен для проверки того, что частная производная $\partial f/\partial y$ правильна, а путь вверх–вправо удобен для $\partial f/\partial x$.

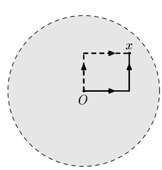


Рис. 5. Пути разные, результат один.

Теперь всё готово для доказательства критерия замкнутости *дифференци-*

руемой формы, определённой в открытом подмножестве плоскости: *форма*

$$P(x,y)\,dx + Q(x,y)\,dy$$

замкнута, если и только если

$$\frac{\partial P}{\partial y}(x,y) = \frac{\partial Q}{\partial x}(x,y)$$

в любой точке (x,y) *её области определения.*

В одну сторону это утверждение о симметрии второй производной: если $P = \partial f/\partial x$ и $Q = \partial f/\partial y$ для локальной первообразной f, то по предположению о дифференцируемости P и Q функция f оказывается дважды дифференцируемой и $\partial^2 f/\partial x\,\partial y = \partial^2 f/\partial y\,\partial x$.

Чтобы доказать обратное утверждение, достаточно проверить, что интеграл по любому прямоугольнику, целиком содержащемуся в области определения формы $P\,dx + Q\,dy$, равен нулю. (Тогда в любом круге она будет точной, и потому замкнута.) Пусть это не так и интеграл по некоторому прямоугольнику превосходит (по модулю) ε, умноженный на площадь этого прямоугольника. Фиксируем такое (положительное) значение ε и назовём прямоугольники с указанным свойством «плохими».

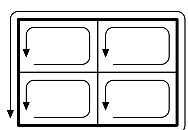


Рис. 6. Одна из частей плохого прямоугольника плоха.

Если прямоугольник разрезан на четвертушки (рис. 6), то интеграл по всему прямоугольнику равен сумме интегралов по частям (внутренние отрезки проходятся туда – обратно и сокращаются). То же самое верно и для площади, поэтому если все части хорошие, то и весь прямоугольник хорош.

Таким образом, у плохого прямоугольника есть плохая четвертушка, её снова разрежем на части и так далее. Получатся вложенные подобные плохие прямоугольники. Их общая точка (α,β) лежит в области определения формы и обладает таким свойством: существует сколь угодно малый охватывающий её плохой прямоугольник с фиксированным отношением сторон.

Это противоречит равенству $\partial P/\partial y(\alpha,\beta) = \partial Q/\partial x(\alpha,\beta)$. В самом деле, рассмотрим прямоугольник $[a,b]\times[c,d]$, содержащий (α,β). Интеграл по нему состоит из интегралов по горизонтальным и вертикальным сторонам и

равен

$$\int_a^b [P(x,c) - P(x,d)]\,dx + \int_c^d [Q(b,y) - Q(a,y)]\,dy$$

(на горизонтальных сторонах важно P, на вертикальных — Q). Используя дифференцируемость P и Q в точке (α,β), можно заменить все четыре величины под интегралом на их линейные приближения, при этом ошибка в каждой из них будет $o(l)$, где l — линейные размеры прямоугольника. Получится

$$\int_a^b \frac{\partial P}{\partial y}(\alpha,\beta)\cdot(c-d)\,dx + \int_c^d \frac{\partial Q}{\partial x}(\alpha,\beta)\cdot(b-a)\,dy =$$
$$= \left[\frac{\partial Q}{\partial x}(\alpha,\beta) - \frac{\partial P}{\partial y}(\alpha,\beta)\right](d-c)(b-a) = 0,$$

при этом ошибка линеаризации под интегралом есть $o(l)$, а с учётом отрезка интегрирования $o(l^2)$, и потому не может быть больше $\varepsilon\cdot$(площадь) для стремящихся к нулю прямоугольников (с постоянным отношением сторон). Полученное противоречие показывает, что плохих прямоугольников не существует, и завершает доказательство.

52 Пусть $f\colon \mathbb{C}\to\mathbb{C}$ — дифференцируемая в комплексном смысле функция. (Это значит, что она дифференцируема как функция из $\mathbb{R}^2$ в $\mathbb{R}^2$ и её производная есть комплексно-линейный оператор, то есть умножение на некоторое комплексное число. Эквивалентное условие: при любом z существует предел $[f(z+h)-f(z)]/h$, когда комплексное h стремится к нулю.) Дайте определение формы $f(z)\,dz$ и покажите, что она точна.

53 (Для знакомых с интегрированием в $\mathbb{R}^2$.) Покажите, что интеграл формы $P(x,y)\,dx + Q(x,y)\,dy$ с *непрерывно дифференцируемыми* коэффициентами по границе прямоугольника, содержащегося в её области определения, равен интегралу

$$\iint \left[\frac{\partial Q}{\partial x}(x,y) - \frac{\partial P}{\partial y}(x,y)\right] dx\,dy$$

по этому прямоугольнику (*формула Грина для прямоугольника*). Это даёт другое доказательство критерия замкнутости для форм с непрерывно дифференцируемыми коэффициентами.

Аналогичный критерий замкнутости есть и для формы с дифференцируемыми коэффициентами, определённой в открытом подмножестве $U\subset\mathbb{R}^n$: *форма*

$$P_1(x_1,\ldots,x_n)\,dx_1 + \ldots + P_n(x_1,\ldots,x_n)\,dx_n$$

замкнута тогда и только тогда, когда $\partial P_i/\partial x_j = \partial P_j/\partial x_i$ во всех точках U.

Доказательство аналогично двумерному случаю. Сначала мы доказываем, что если форма определена в шаре и интеграл по любому двумерному прямоугольнику (все координаты, кроме двух, постоянны, а две переменные координаты описывают прямоугольник) равен нулю, то она точна в этом шаре. В самом деле, от любого из $n!$ ломаных путей, ведущих из центра шара в x (сначала по одной координате, потом по другой и так далее) можно перейти к любому другому с помощью транспозиций, которые соответствуют изменению обхода прямоугольника. Поэтому интегралы по всем этим $n!$ путям одинаковы. Мы можем выбирать те из них, которые нам подходят, и рассуждать как раньше.

А интеграл по двумерному прямоугольнику (в плоскости (x_i, x_j)) зависит только от P_i и P_j, так что можно применить доказанное нами двумерное утверждение.

21. Гомотопные пути

Замкнутые формы локально точны, поэтому можно ожидать, что интеграл такой формы не меняется при небольшом изменении пути (а следовательно, и при любом постепенном изменении, поскольку оно разбивается на небольшие). Точная формулировка этого утверждения использует понятие гомотопных путей.

Пусть $\gamma_0, \gamma_1 : [0,1] \to X$ — два пути в топологическом пространстве X с общими началом и концом, то есть $\gamma_0(0) = \gamma_1(0)$ и $\gamma_0(1) = \gamma_1(1)$. Говорят, что они *гомотопны как пути с фиксированным началом и концом*, если их можно включить в семейство путей $\gamma_t : [0,1] \to X$ с тем же началом и концом, где параметр t меняется от нуля до единицы, причём отображение

$$(t, s) \mapsto \gamma_t(s)$$

является непрерывным отображением квадрата $[0,1] \times [0,1]$ в X. Другими словами, пути гомотопны, если непрерывное отображение границы квадрата в пространство X, которое слева совпадает с γ_0, справа — с γ_1, а на верхней и нижних сторонах постоянно, продолжается до отображения всего квадрата (с внутренностью), см. рис. 7.

54 Дайте определение гомотопных петель (*петля* — путь, у которого начало и конец совпадают). Проверьте, что пути γ_0 и γ_1 с общим началом и концом гомотопны (в смысле данного выше определения) тогда и только тогда, когда петля «туда по γ_0, обратно по γ_1» гомотопна постоянной.

Теперь можно сформулировать обещанное утверждение: *интегралы замкнутой формы по гомотопным путям равны.* (Рассматривается форма, опре-

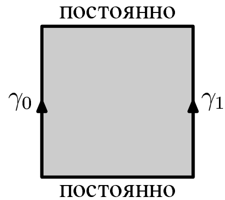


Рис. 7. Гомотопия.

деленная на некотором многообразии X, и кусочно-гладкие пути, гомотопные в X.)

Чтобы доказать это, посмотрим на прямоугольник гомотопии (рис. 7). Когда точка обходит границу этого прямоугольника против часовой стрелки, её образ движется по X вдоль γ_1 и обратно по γ_0, а в промежутках стоит. Образуется петля в X, и нам надо доказать, что интеграл по этой петле равен нулю.

Разрежем квадрат на много мелких, и рассмотрим петли, обходящие каждый из получившихся при этом квадратиков (рис. 8). Как и раньше, инте-

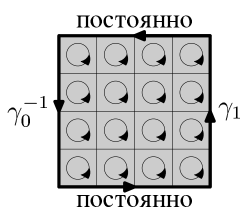


Рис. 8. Разложение в сумму интегралов по малым петлям.

грал по большой петле равен сумме интегралов по малым. (Тут есть трудность: малые петли лишь непрерывны, а не кусочно-гладкие, так что интеграл по ним не определён — но об этом после, пусть пока эти петли оказались гладкими.)

Если разбиение настолько мелкое, что образ каждого квадратик целиком попадает в область, где форма точна, то интегралы по малым петлям равны нулю (как и положено для точных форм, см. выше). Поэтому и весь интеграл равен нулю.

Чтобы найти такое мелкое разбиение, для каждой точки u квадрата рассмотрим её образ v в X; по определению замкнутой формы найдётся окрестность точки v, в которой форма точна. Прообраз этой окрестности (при

непрерывном отображении квадрата, задающем гомотопию) будет окрестностью точки u. Получаем покрытие компакта (квадрата) открытыми множествами. Остаётся воспользоваться леммой Лебега, согласно которой существует такое ε, что любой шар радиуса ε целиком попадает в одно из множеств покрытия, и взять сторону квадратиков меньшей $\varepsilon/2$.

Если внутренние пути не гладкие, мы можем попросту *определить* интеграл как разницу значений соответствующей первообразной (которая существует по доказанному) в конце и начале пути. При этом равенство нулю интеграла по границе квадратика будет очевидно. Чуть менее очевидно, что при сложении интегралов по границам квадратиков получится исходный интеграл, то есть что внутренние пути сократятся. Проблема в том, что при вычислении интегралов на пути туда и обратно мы используем разные первообразные (для соседних квадратиков). Но их разность постоянна на любом связном множестве, в частности, на образе пути, так что при вычислении интеграла обе первообразные дадут один и тот же результат.

55 Докажите, что интеграл от замкнутой формы по кусочно-гладкой петле, гомотопной постоянной, равен нулю.

56 Дайте определение интеграла от замкнутой формы по любому непрерывному (не обязательно гладкому) пути и докажите, что он не меняется при гомотопии.

Доказанная теорема позволяет строго определить угол поворота кусочно-гладкого пути (или непрерывного, см. предыдущую задачу) $\gamma\colon [0,1] \to \mathbb{R}^2 \setminus \{(0,0)\}$ как интеграл от формы $d\varphi$ по этому пути. Это понятие имеет многочисленные топологические применения (основная теорема алгебры, теорема о неподвижной точке для круга, пересечение любых двух непрерывных путей, соединяющих противоположные вершины квадрата внутри него и др.)

57 Докажите, что для замкнутого пути, не проходящего через начало координат, интеграл от формы $d\varphi$ кратен 2π.

22. Векторные поля

Пусть имеется некоторое многообразие X и в каждой точке $x \in X$ выбран некоторый элемент $v(x)$ касательного пространства $T_x(X)$, то есть касательный вектор в точке X. В этом случае говорят, что на многообразии X задано *векторное поле* v.

В карте $\varphi\colon U \subset \mathbb{R}^k \to X$ возникает изображение векторного поля: для каждой точки $u \in U$ можно рассмотреть набор чисел, являющихся коорди-

натами вектора $v(\varphi(u))$ в этой карте, то есть элемент $\mathbb{R}^k$, соответствующий $v(\varphi(u))$ при изоморфизме пространств $\mathbb{R}^k$ и $T_{\varphi(u)}(X)$.

Векторное поле называется *непрерывным* (*дифференцируемым*, *гладким*), если таковы его изображения во всех картах. Это не зависит от выбора карт при достаточной гладкости функций перехода (так что понятие непрерывного векторного поля имеет смысл для C^1-многообразий, дифференцируемого — для C^2-многообразий и т. д.).

58 Построить непрерывное векторное поле на окружности, которое ни в одной точке не обращается в нуль. Тот же вопрос для тора и для листа Мёбиуса.

Оказывается, что на двумерной сфере непрерывного и всюду ненулевого векторного поля не существует — «ежа нельзя причесать». Эта теорема доказывается в курсах топологии (одно из доказательств использует понятие числа оборотов вокруг точки, о котором мы говорили).

59 Покажите, что на трёхмерной сфере $S^3 \subset \mathbb{R}^4$ существует непрерывное ненулевое векторное поле и даже три таких поля, в каждой точке образующих базис касательного пространства. [Указание. Полезно использовать умножение кватернионов.]

Дифференцируемый путь $\gamma\colon (a, b) \to X$ на многообразии X называется *траекторией* векторного поля v, если

$$\gamma'(t) = v(\gamma(t))$$

для любого $t \in (a, b)$.

Если дорожная полиция поставила в каждой точке многообразия X указатель, говорящий, с какой скоростью и куда надо ехать, то траектория — это поведение законопослушного водителя. При этом оказывается, что указания всегда выполнимы и однозначно определяют движение (при данной начальной точке). А именно, *для данного непрерывно дифференцируемого векторного поля v на многообразии X, данной точки $x_0 \in X$ и данного момента времени $t_0 \in \mathbb{R}$ всегда найдётся $\delta > 0$ и траектория*

$$\gamma\colon (t_0 - \delta, t_0 + \delta) \to X,$$

для которой $\gamma(t_0) = x_0$. (Заметим, что траектория может существовать не для всякого δ — например, $v(x)$ может так быстро удлиняться, что траектория вылетает с многообразия за конечное время.)

Кроме того, траектория единственна: *если две траектории $\gamma_1\colon (a_1, b_1) \to X$ и $\gamma_2\colon (a_2, b_2) \to X$ проходят через одну и ту же точку в один и тот же момент времени, то есть $\gamma_1(t_0) = \gamma_2(t_0)$ для*

некоторого $t_0 \in (a_1, b_1) \cap (a_2, b_2)$*, то они совпадают на общей области определения:* $\gamma_1(t) = \gamma_2(t)$ *при всех* $t \in (a_1, b_1) \cap (a_2, b_2)$.

Существование и единственность траекторий есть частный случай общих утверждений о дифференциальных уравнениях, к которым мы и переходим.

23. Дифференциальные уравнения

Пусть дано банахово пространство Y, открытое множество $U \subset \mathbb{R} \times Y$ и функция $F\colon U \to Y$. Тогда имеет смысл говорить о *дифференциальном уравнении*

$$y'(t) = F(t, y(t)),$$

решениями которого могут быть (или не быть) дифференцируемые функции

$$y\colon (a, b) \to Y,$$

определённые на некотором интервале действительной прямой. (При этом, естественно, требуется, чтобы точка $(t, y(t))$ попадала в U при всех $t \in (a, b)$.)

60 Покажите, что если функция F гладкая, то любое решение дифференциального уравнения $y'(t) = F(t, y(t))$ является гладкой функцией.

При определённых условиях на F можно гарантировать существование и единственность решения с данным *начальным условием* $(t_0, y_0) \in U$ (это значит, что мы дополнительно требуем $y(t_0) = y_0$) — по крайней мере локально, в окрестности точки t_0.

Более подробно. Пусть функция F непрерывна и удовлетворяет условию Липшица по Y-координате с некоторой константой L:

$$\|F(x, y_1) - F(x, y_2)\| \leqslant L \cdot \|y_1 - y_2\|$$

(норма в обеих частях берётся в пространстве Y). Тогда *для любого начального условия* (t_0, y_0) *существует такое* $\delta > 0$*, что решение дифференциального уравнения* $y'(t) = F(t, y(t))$ *на интервале* $(t_0 - \delta, t_0 + \delta)$ *с начальным условием* $y(t_0) = y_0$ *существует и единственно.*

Заметим, что если функция F непрерывно дифференцируема, то её производная ограничена в некоторой окрестности точки (t_0, y_0), и в силу теоремы о конечных приращениях в этой окрестности выполняется условие Липшица (в качестве L надо взять верхнюю границу для нормы производной).

Докажем существование. Дифференциальное уравнение и начальное условие соединим в интегральное уравнение

$$y(t) = y_0 + \int_{t_0}^{t} F(\xi, y(\xi))\, d\xi.$$

61 Проверьте, что это уравнение (на непрерывную функцию $y\colon (a, b) \to Y$) действительно равносильно дифференциальному уравнению вместе с начальным условием.

Это интегральное уравнение можно трактовать так: функция y является неподвижной точкой отображения Φ, которое отображает y в функцию $\Phi(y)$, для которой

$$\Phi(y)\colon\, t \mapsto y_0 + \int_{t_0}^{t} F(\xi, y(\xi))\, d\xi.$$

Поэтому достаточно подобрать полное метрическое пространство, в котором этот оператор будет сжимающим. В качестве такого пространства мы рассмотрим пространство непрерывных отображений интервала $(t_0 - \delta, t_0 + \delta)$ в замкнутый шар $B_\varepsilon(y_0)$ (шар радиуса ε с центром в y_0 в пространстве Y) при подходящих δ и ε. Это пространство рассматривается с метрикой равномерной сходимости.

От δ и ε потребуем, чтобы

• $(t_0 - \delta, t_0 + \delta) \times B_\varepsilon(y_0) \subset U$ (иначе правая часть уравнения может не иметь смысла);

• функция F была ограничена в $(t_0-\delta, t_0+\delta)\times B_\varepsilon(y_0)$ некоторой константой C, причём $C\delta \leqslant \varepsilon$ (это гарантирует, что значение $[\Phi(y)](t)$ не выйдет из шара $B_\varepsilon(y_0)$, поскольку подинтегральное [или подынтегральное, если буквально следовать правилам русского языка] выражение не превосходит C по норме);

• $L\delta \leqslant 1/2$.

Последнее условие нужно, чтобы оператор Φ был сжимающим с коэффициентом $1/2$. В самом деле, тогда

$$\begin{aligned}
\|\Phi(y_1) - \Phi(y_2)\| &= \sup_{t\in(t_0-\delta, t_0+\delta)} \|[\Phi(y_1)](t) - [\Phi(y_2)](t)\| = \\
&= \sup_{t\in(t_0-\delta, t_0+\delta)} \left\| \int_{t_0}^{t} [F(\xi, y_1(\xi)) - F(\xi, y_2(\xi))]\, d\xi \right\| \leqslant \\
&\leqslant \sup_{t\in(t_0-\delta, t_0+\delta)} \int_{t_0}^{t} \|F(\xi, y_1(\xi)) - F(\xi, y_2(\xi))\|\, d\xi \leqslant \\
&\leqslant \sup_{t\in(t_0-\delta, t_0+\delta)} \int_{t_0}^{t} L\cdot \|y_1(\xi) - y_2(\xi)\|\, d\xi \leqslant \delta\cdot L\cdot \|y_1 - y_2\| \leqslant \tfrac{1}{2}\|y_1 - y_2\|.
\end{aligned}$$

(эти равенства написаны для $t > t_0$, при $t < t_0$ надо переставить пределы интегрирования в двух последних строках). Таким образом, мы имеем сжимающее отображение замкнутого шара в полном метрическом пространстве (ограниченных непрерывных функций $(t_0 - \delta, t_0 + \delta) \to Y$) в себя, и потому существует неподвижная точка.

Осталось сказать, в каком порядке мы выбираем δ и ε: сначала мы выбираем их достаточно малыми, чтобы выполнялось первое условие и чтобы F было ограничено некоторой константой C, а затем уменьшаем δ так, чтобы выполнялись требуемые неравенства.

Существование доказано. Единственность доказана для функций, не выходящих из шара $B_\varepsilon(y_0)$, и остаётся заметить, что при наших предположениях на δ и ε решение не может выйти из этого шара. В самом деле, пусть оно в первый раз выходит на границу шара в момент времени $t_0 + \sigma$, где σ — некоторое положительное число, меньшее δ. Тогда на $(t_0, t_0 + \sigma)$ производная $y'(t) = F(t, y(t))$ по норме не превосходит C, поэтому получаем противоречие с теоремой о конечных приращениях и неравенством $C\delta \leqslant \varepsilon$. Случай $\sigma < 0$ аналогичен.

Замечание. Поскольку мы можем произвольно уменьшать δ, мы доказали также и такое утверждение: любые два решения, определённые на (сколько угодно малых) интервалах, содержащих t_0, совпадают в некоторой окрестности t_0. В самом деле, можно уменьшить δ так, чтобы интервал $(t_0 - \delta, t_0 + \delta)$ попал в область определения обоих решений, и воспользоваться доказанным утверждением о единственности.

62 Рассмотрим уравнение $y'(t) = F(t, y(t))$ с непрерывной правой частью, локально удовлетворяющей условию Липшица. Докажите, что любые два решения с начальным условием $y(t_0) = y_0$, определённые на произвольных интервалах, содержащих t_0, совпадают на пересечении областей определения. [Указание: рассмотрим точную верхнюю и нижнюю грань тех моментов времени, до которых они совпадают.]

63 Докажите, что для уравнения с непрерывной правой частью решение может не быть единственным, построив уравнение, которое имеет решения $y(t) = 0$ и $y(t) = \textbf{if } t \leqslant 0 \textbf{ then } 0 \textbf{ else } t^2 \textbf{ fi}$.

64 Докажите теорему о (локальном) существовании решения уравнения с непрерывной правой частью. Для этого рассмотрите *ломаные Эйлера*, в которых ось времени разбивается на отрезки равной длины и скорость на каждом отрезке постоянна и определяется дифференциальным уравнением в его начале. Покажите, что из последовательности ломаных Эйлера со стремящейся к нулю длиной отрезка можно выбрать сходящуюся подпоследовательность. (По лемме Арцела – Асколи достаточно доказать, что они

равномерно ограничены и равностепенно равномерно непрерывны, и это несложно.) Докажите, что предел этой подпоследовательности удовлетворяет интегральному (а значит, и дифференциальному) уравнению.

24. Сдвиг вдоль векторного поля

Вернёмся к траекториям векторных полей. Локально существование и единственность можно доказывать в картах, где это сводится к уже доказанным теоремам о дифференциальных уравнениях (даже к их частному случаю, когда уравнение *автономно*, то есть его правая часть $F(t, y)$ не зависит от t.)

Отсюда можно вывести и глобальную единственность траектории, о которой мы говорили. В самом деле, рассмотрим точную верхнюю грань τ тех моментов времени, до которых траектории совпадают

$$\tau = \sup\{t \mid \gamma_1(\xi) = \gamma_2(\xi) \text{ при всех } \xi \in (t_0, t)\}.$$

Если τ выходит из области определения хотя бы одной из траекторий, то доказывать нечего. Если в τ определены обе траектории, то $y_1(\tau) = y_2(\tau)$ в силу непрерывности и — внимание! — отделимости многообразия. Тогда локальная теорема о единственности показывает, что траектории совпадают и в некоторой окрестности момента τ, то есть он не был точной верхней гранью.

Таким образом, для данной точки x и данного непрерывно дифференцируемого векторного поля v можно выпустить единственную траекторию этого поля (соединение всех траекторий — в силу единственности они продолжают друг друга). Она может быть определена на всей прямой или на её части (интервале или луче). Выбор начального момента времени не играет роли (с точностью до сдвига), поэтому можно говорить о *сдвиге точки* $x \in X$ *вдоль векторного поля* v *на время* t. Надо только иметь в виду, что сдвиг определён не при всех x и t (мы можем «выйти за пределы многообразия»).

65 Докажите, что если многообразие компактно, то возможен сдвиг любой точки на любое время. [Указание. Что может случиться с траекторией, если многообразие компактно? она должна иметь некоторую предельную точку, и если она появляется в окрестности этой предельной точки, то далеко не уйдёт, поскольку векторное поле ограничено.]

66 Докажите аналогичное утверждение для *финитного* векторного поля (=равного нулю вне некоторого компактного подмножества многообразия).

Для произвольного (не обязательно компактного) многообразия X с векторным полем v функция

$$(x, t) \mapsto \text{сдвиг точки } x \text{ вдоль } v \text{ на } t$$

определена на некотором подмножестве $X \times \mathbb{R}$ (и принимает значения в X). Можно доказать, что если многообразие и векторное поле гладкие, то область определения этой функции открыта, а функция является гладкой.

(Это следует из теорем о гладкой зависимости решений дифференциальных уравнений от параметров. Их, в свою очередь, можно вывести из теоремы о неявной функции, если рассматривать неподвижность точки в доказательстве теоремы существования как уравнение, которое мы разрешаем относительно неизвестной функции y. Мы не будем про это говорить подробно и примем утверждение о гладкости указанной функции без доказательства.)

Оказывается, что в окрестности неособой точки гладкого поля (то есть точки, в которой поле не равно нулю) любое поле устроено как постоянное. *Пусть X — гладкое многообразие, v — гладкое векторное поле на X и $x \in X$ — неособая точка этого поля, то есть $v(x) \neq 0$. Тогда существует карта, покрывающая точку x, в которой это векторное поле изображается постоянным векторным полем.*

Ясно, что многообразия тут не по существу, поскольку теорема всё равно локальна, и можно рассматривать поле в открытом множестве $U \subset \mathbb{R}^k$. Для упрощения обозначений считаем, что $x = (0, \ldots, 0)$. Рассмотрим вектор $v(0)$ и выберем в $\mathbb{R}^k$ подпространство Γ размерности $k-1$, не содержащее этого вектора и потому дающее в прямой сумме с ним всё $\mathbb{R}^k$ (рис. 9).

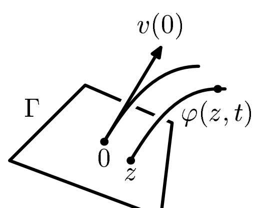


Рис. 9. Выпрямление векторного поля.

Рассмотрим теперь отображение $\varphi\colon \Gamma \times \mathbb{R} \to \mathbb{R}^k$, положив $\varphi(z, t)$ равным сдвигу точки $z \in \Gamma$ на t вдоль поля v. К этому гладкому отображению можно применить теорему об обратной функции, поскольку его производная в точке $(0, 0)$ обратима: частная производная по $z \in \Gamma$ является единичным

оператором, так как $\varphi(z,0) = z$, частная производная по $t \in \mathbb{R}$ равна $v(0)$, и $v(0)$ по построению не лежит в Γ. Таким образом, в окрестности нуля отображение φ является диффеоморфизмом, и при этом диффеоморфизме траектории поля переходят в прямые в пространстве $\Gamma \times \mathbb{R}$, параллельные второй координате и проходимые с единичной скоростью. А значит, векторное поле переходит при этом диффеоморфизме в постоянное.

Замечание. Теорема о выпрямлении векторного поля существенно локальна — хотя бы потому, что у поля могут быть циклические траектории.

Таким образом, мы видим, что в неособых точках все векторные поля (с точностью до диффеоморфизма) устроены одинаково. Заметим, что для дифференциальных форм первой степени (ковекторных полей) это было совсем не так — некоторые из них замкнуты, а некоторые — нет, и это свойство сохраняется при диффеоморфизме.

25. Коммутатор векторных полей

Пусть теперь на многообразии X есть не одно векторное поле, а два поля v и w. Можно ли их выпрямить одновременно (найти карту, в которой v и w оба изображаются постоянными полями)?

Необходимое условие: сдвиги вдоль них должны коммутировать. Это означает следующее. Обозначим через v_t отображение многообразия X в себя, сдвигающее каждую точку вдоль v на t; аналогично для w. (Эти отображения не обязательно определены на всём X, особенно при бо́льших t, но локально — в окрестности данной точки и при малом t — они определены.) Так вот, если два поля допускают совместное выпрямление в окрестности некоторой точки a, то

$$v_t(w_s(x)) = w_s(v_t(x))$$

при всех x из некоторой окрестности точки a и при всех достаточно малых s и t. (В самом деле, после выпрямления сдвиги коммутируют.)

67 Рассмотрим на земной сфере два векторных поля из векторов единичной длины: одно направлено на север, другое — на восток. Можно ли их одновременно выпрямить в окрестности Москвы? Как надо изменить длину векторов восточного поля, чтобы поля допускали одновременное выпрямление?

Как и в случае замкнутых дифференциальных форм, хотелось бы получить явный критерий коммутирования сдвигов. Чтобы сделать это, определим операцию *дифференцирования гладкой функции* $\varphi\colon X \to \mathbb{R}$ *вдоль векторного поля* u. Её результатом является другая гладкая функция $X \to \mathbb{R}$,

которую мы будем обозначать $v\varphi$ и которая определяется формулой

$$(v\varphi)(x) = \varphi'(x)v(x)$$

(в правой части производная $\varphi'(x)$, являющаяся линейным отображением $T_x(X)$ в $\mathbb{R}$, применяется к вектору $v(x) \in T_x(X)$).

68 Покажите, что для произвольной гладкой функции φ и точки x производная функции $t \mapsto \varphi(v_t(x))$ в точке t равна $(v\varphi)(v_t(x))$.

Для данного (гладкого) векторного поля v на многообразии X получается линейное отображение $\varphi \mapsto v\varphi$ алгебры гладких функций $C^\infty(X)$ в себя, которое является *дифференцированием*, то есть удовлетворяет *тождеству Лейбница*:

$$v(\varphi \cdot \psi) = (v\varphi) \cdot \psi + \varphi \cdot (v\psi)$$

69 (**а**) Проверьте это. (**б**) Покажите, что всякое дифференцирование D алгебры $C^\infty(X)$ задаётся некоторым гладким векторным полем на X. [Указание. Для начала надо проверить, что если функция φ равна нулю в окрестности некоторой точки x, то и $D\varphi$ равно нулю в x. Это можно доказать, представив φ в виде произведения двух функций, равных нулю в x. После этого ясно, что $D\varphi(x)$ определяется ростком функции φ в точке x, и можно применять лемму Адамара.]

Теперь можно сформулировать необходимое условие коммутирования сдвигов в дифференциальной форме: *если сдвиги вдоль векторных полей v и w коммутируют, то*

$$v(w\varphi) - w(v\varphi) = 0$$

для любой гладкой функции φ, то есть коммутируют соответствующие дифференцирования.

В самом деле, рассмотрим (для данной точки x) гладкую функцию двух действительных переменных

$$(s, t) \mapsto \varphi(v_t(w_s(x))).$$

Её частная производная по t при произвольных s и t равна $(v\varphi)(v_t(w_s(x))$; при $t = 0$ получается $(v\varphi)(w_s(x))$, поэтому второе дифференцирование по s даёт в точке $(0, 0)$ смешанную производную $w(v\varphi)(x)$. Если $v_t \circ w_s = w_s \circ v_t$ при всех достаточно малых s и t, то симметрия второй производной даёт $w(v\varphi) = v(w\varphi)$, как и утверждалось.

Оказывается, что коммутатору дифференцирований вдоль v и w соответствует дифференцирование вдоль некоторого третьего векторного поля:

для любых двух гладких векторных полей v *и* w *на многообразии* X *существует и единственно векторное поле, называемое их коммутатором и обозначаемое* $[v, w]$, *для которого*

$$[v, w]\varphi = v(w\varphi) - w(v\varphi).$$

Единственность очевидна (всякое поле определяется результатами его применения к координатным функциям). Если бы мы знали, что всякое дифференцирование алгебры гладких функций задаёт векторное поле (см. задачу 69), то достаточно было бы проверить, что правая часть является дифференцированием, что делается чисто формально:

$$\begin{aligned}
[v, w](\varphi\psi) = v(w(\varphi\psi)) - w(v(\varphi\psi)) = \\
= v(w(\varphi)\psi) + v(\varphi w(\psi)) - w(v(\varphi)\psi) - w(\varphi v(\psi)) = \\
= v(w(\varphi))\psi + w(\varphi)v(\psi) + v(\varphi)w(\psi) + \varphi v(w(\psi)) - \\
- w(v(\varphi))\psi - v(\varphi)w(\psi) - w(\varphi)v(\psi) - \varphi w(v(\psi)) = \\
= v(w(\varphi))\psi - w(v(\varphi))\psi + v(w(\psi))\varphi - w(v(\psi))\varphi = \\
= [v, w]\varphi \cdot \psi + \varphi \cdot [v, w]\psi.
\end{aligned}$$

Можно доказать существование и другим способом, явно вычислив координаты поля $[v, w]$ в некоторой карте. Если поле v изображается в этой карте как $(v^1, \ldots, v^n)$ (где каждое v^i есть функция от n действительных аргументов $(x_1, \ldots, x_n)$ с действительными значениями), а w изображается как $(w^1, \ldots, w^n)$, то

$$w\varphi = \sum_j w^j \frac{\partial\varphi}{\partial x_j}$$

и

$$v(w\varphi) = \sum_{i,j} v^i \frac{\partial}{\partial x_i}\left(w^j \frac{\partial\varphi}{\partial x_j}\right) = \sum_{i,j} v^i w^j \frac{\partial^2\varphi}{\partial x_i\,\partial x_j} + \sum_{i,j} v^i \frac{\partial w^j}{\partial x_i}\frac{\partial\varphi}{\partial x_j}$$

При вычитании аналогичного выражения с переставленными v и w первый член сокращается (симметрия второй производной), и всего получается

$$v(w\varphi) - w(v\varphi) = \sum_j \left(\sum_i v^i \frac{\partial w^j}{\partial x_i} - \sum_i w^i \frac{\partial v^j}{\partial x_i}\right)\frac{\partial\varphi}{\partial x_j}.$$

то есть коммутатор дифференцирований вдоль v и w есть (в карте) дифференцирование вдоль векторного поля с координатами

$$[v, w]^j = \sum_i v^i \frac{\partial w^j}{\partial x_i} - \sum_i w^i \frac{\partial v^j}{\partial x_i} = [w'v - v'w]^j \qquad (*)$$

так что $[v, w] = w'v - v'w$ (здесь w' и v' — производные векторных полей w и v, рассматриваемых как отображения $\mathbb{R}^n \to \mathbb{R}^n$; в каждой точке эти производные суть линейные операторы, которые применяются к векторам v и w).

Формулу $(*)$ можно принять за определение коммутатора векторных полей и не говорить ничего о дифференцированиях, но только тогда надо проверять, что задаваемый ею набор чисел преобразуется при переходе из карты в карту так, как подобает вектору.

Геометрический смысл коммутатора можно описать так. Пусть φ — произвольная гладкая функция, определённая в окрестности точки x. Рассмотрим функцию

$$(s, t) \mapsto B(s, t) = \varphi(v_t(w_s(x))) - \varphi(w_s(v_t(x))$$

Очевидно, что $B(0, t) = B(s, 0) = 0$, поэтому частные производные $\partial B/\partial s(0, 0)$ и $\partial B/\partial t(0, 0)$, а также вторые частные производные $\partial^2 B/\partial s^2(0, 0)$ и $\partial^2 B/\partial t^2(0, 0)$ равны нулю. Смешанную производную мы по существу уже вычисляли (отдельно для каждого из членов) и знаем, что

$$\frac{\partial^2 B}{\partial s\, \partial t} = [v(w\varphi) - w(v\varphi)](x) = ([v, w]\varphi)(x).$$

Поэтому формула Тейлора с точностью до членов второго порядка даёт

$$B(s, t) = \varphi(v_t(w_s(x))) - \varphi(w_s(v_t(x)) = ([v, w]\varphi)(x) \cdot st + o((|s| + |t|)^2).$$

Взяв в качестве φ координатные функции в некоторой карте, мы получаем, что в этой карте

$$v_t(w_s(x)) - w_s(v_t(x) = [v, w] \cdot st + o((|s| + |t|)^2).$$

Таким образом, главный член разницы между сдвигами точки в разных порядках (с точностью до второго порядка) пропорционален st, а коэффициент пропорциональности и есть коммутатор (рис. 10).

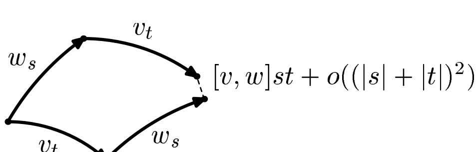


Рис. 10. Геометрический смысл коммутатора.

Заметим, что такого рода сравнения возможны только в карте (иначе слева стояла бы разность между двумя точками многообразия, не имеющая смысла, а справа — касательный вектор в третьей точке).

Оказывается, что условие $[v, w] = 0$ не только необходимо, но и достаточно для того, чтобы сдвиги вдоль v и w коммутировали. *Пусть векторные поля v и w определены в окрестности точки x на многообразии X и $[v, w] = 0$ во всех точках этой окрестности. Тогда при достаточно малых s и t сдвиги $v_s w_t x$ и $w_t v_s x$ определены и равны.*

Слова «на многообразии» здесь не по существу: утверждение локально, так что его можно проверять в карте и считать, что векторные поля заданы в некоторой окрестности точки $x \in \mathbb{R}^n$.

Наиболее естественное (хотя и требующее дополнительных уточнений) рассуждение состоит в следующем. Пусть коммутатор равен нулю в окрестности точки x. Чтобы показать, что $v_t w_s x - w_s v_t x$ равно нулю (а не только $o((|s| + |t|)^2)$, что мы уже знаем), разобьём прямоугольник $[0, t] \times [0, s]$ на n^2 маленьких и будем идти из угла $A(0, 0)$ в $C(s, t)$ различными способами по сторонам маленьких прямоугольников (рис. 11). Каждому из путей в

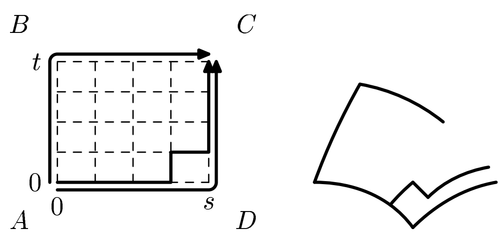


Рис. 11. Переход за n^2 шагов.

прямоугольнике соответствует последовательность сдвигов вдоль векторных полей; от пути ADC можно перейти к ABC за n^2 шагов; каждый переход огибает один маленький прямоугольник (меняя способ его обхода). Как будет сдвигаться конец криволинейного пути? При обходе с разных сторон одного прямоугольника возникнет расхождение порядка $o(1/n^2)$, которое при дальнейшем движении может ещё увеличиться. Если поверить, что это увеличение происходит не более чем в константу раз, а также что все $o(1/n^2)$ оцениваются равномерно, то видно, что расстояние между образами путей ADC и ABC (точнее, их концами) не превосходит $n^2 O(1) o(1/n^2) = o(1)$. Устремляя n к бесконечности, заключаем, что это расстояние равно нулю.

70 Проведите аккуратно необходимые оценки.

В окрестности точки, где одно из полей (скажем, v) не равно нулю,

есть более простой способ доказательства. Теорема о выпрямлении позволяет считать, что это поле постоянно. Тогда

$$[v, w] = w'v - v'w = w'v - 0 = w'v,$$

и потому $[v, w] = 0$ означает, что поле w постоянно вдоль траекторий (прямых) поля v. Поэтому сдвиг траектории поля w на любой вектор, параллельный v, переводит её в другую траекторию того же поля, и поля коммутируют.

Ещё одно доказательство использует ту же идею, но обходится без выпрямления. Докажем, что диффеоморфизм v_t оставляет на месте поле w. Тогда применение v_t к траекториям поля w снова даёт траектории поля w, и сдвиги коммутируют. Надо доказать, таким образом, что

$$v_t'(y)w(y) = w(v_t(y))$$

для любой точки y (в окрестности интересующей нас точки x) и для любого достаточно малого t. Перепишем это равенство как

$$w(y) = [v_t'(y)]^{-1}w(v_t(y));$$

нам надо доказать, что правая часть не зависит от t, для чего достаточно продифференцировать её по t и убедиться, что производная равна нулю. В ней переменный оператор $[v_t'(y)]^{-1}$ применяется к переменному вектору $w(v_t(y))$; операция применения билинейна, так что можно дифференцировать по правилу Лейбница. Оператор получается применением обращения $(A \mapsto A^{-1})$ к $v_t'(y)$, поэтому можно воспользоваться теоремой о производной композиции, памятуя, что производная обращения в точке A есть $H \mapsto -A^{-1}HA^{-1}$. Получаем

$$\begin{aligned}\frac{\partial}{\partial t}\left([v_t'(y)]^{-1}w(v_t(y))\right) &= \\ &= [v_t'(y)]^{-1}w'(v_t(y))v(v_t(y)) - [v_t'(y)]^{-1}\frac{\partial}{\partial t}(v_t'(y))[v_t'(y)]^{-1}w(v_t(y)).\end{aligned}$$

Заметим, что $\frac{\partial}{\partial t}(v_t'(y))$ есть смешанная частная производная функции $(t, y) \mapsto v_t(y)$ в точке (t, y) (сначала дифференцируем по y, а потом по t). Переставляя порядок дифференцирования (симметрия) и замечая, что производная $v_t(y)$ по t по определению есть $v(v_t(y))$, получаем

$$\frac{\partial}{\partial t}(v_t'(y)) = \frac{\partial}{\partial y}v(v_t(y)) = v'(v_t(y))v_t'(y).$$

Таким образом, производная оказывается равной

$$[v_t'(y)]^{-1}w'(v_t(y))v(v_t(y)) - [v_t'(y)]^{-1}v'(v_t(y))v_t'(y)[v_t'(y)]^{-1}w(v_t(y)),$$

что после сокращений даёт

$$[v_t'(y)]^{-1}\left(w'(v_t(y))v(v_t(y)) - v'(v_t(y))w(v_t(y))\right) = 0,$$

что и требовалось (в скобке стоит коммутатор в точке $v_t(y)$, равный нулю).

Следствие: *если в окрестности точки $x \in \mathbb{R}^n$ коммутатор векторных полей v и w равен нулю, а в x вектора $v(x)$ и $w(x)$ линейно независимы, то существует диффеоморфизм некоторой окрестности точки x, при котором оба поля переходят в постоянные.*

В самом деле, считаем для простоты записи, что $x = 0$. Возьмём $(n-2)$-мерное подпространство $\Gamma \subset \mathbb{R}^n$, базис которого вместе с $v(x)$ и $w(x)$ образует базис в $\mathbb{R}^n$. Рассмотрим отображение $\Gamma \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n$, задаваемое формулой

$$(z, s, t) \mapsto v_t(w_s(z)) \quad [= w_s(v_t(z))]$$

В точке $(0, 0, 0)$ его частная производная по z есть тождественное отображение Γ в себя, а производные по s и t суть $w(x)$ и $v(x)$, так что можно применить теорему об обратной функции и заключить, что локально это диффеоморфизм. При этом диффеоморфизме траектории обоих полей переходят в прямые, проходимые с постоянной скоростью (для v это ясно из определения, а для w надо воспользоваться переформулировкой в квадратных скобках).

Можно ли утверждать, что при $[v, w] = 0$ отображения v_t и w_s коммутируют при всех (а не только малых) t и s? Тут нужна определённая предосторожность.

71 Приведите пример открытой области $U \subset \mathbb{R}^2$ и двух полей v, w в ней, для которых $[v, w] = 0$ всюду, но при некоторых s, t и x значения $v_t(w_s(x))$ и $w_s(v_t(x))$ определены и не равны. [Указание. См. рис. 12.]

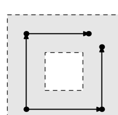


Рис. 12. Локально коммутируют, глобально — нет.

72 Докажите, что если v и w — гладкие векторные поля на компактном многообразии, для которых $[v, w] = 0$ во всех точках многообразия, то v_t и w_s коммутируют при любых t и s. [Указание. Докажите это сначала для произвольных t и малых s.]

ОПЕЧАТКИ, ОШИБКИ, ГЛУПОСТИ,. . .